Fisheries and Oceans Canada | Pêches et Océans Canada
Science | Sciences

C S A S | **S C C S**

Canadian Science Advisory Secretariat | **Secrétariat canadien de consultation scientifique**

Research Document 2003/057 | **Document de recherche 2003/057**

Not to be cited without
Permission of the authors * | Ne pas citer sans
autorisation des auteurs *

Local Influence Diagnostics for the Retrospective Problem in Sequential Population Analysis.

Diagnostics d'influence locale concernant le profil rétrospectif dans l'analyse séquentielle de population

By N. G. Cadigan[1] and P. J. Farrell[2]

[1]Fisheries and Oceans
Northwest Atlantic Fisheries Center
P.O. Box 5667
St. John's, NL
A1C 5X1

[2]School of Mathematics and Statistics
Carleton University
Ottawa, ON
K1S 5B6

* This series documents the scientific basis for the evaluation of fisheries resources in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Research documents are produced in the official language in which they are provided to the Secretariat.

This document is available on the Internet at:

* La présente série documente les bases scientifiques des évaluations des ressources halieutiques du Canada. Elle traite des problèmes courants selon les échéanciers dictés. Les documents qu'elle contient ne doivent pas être considérés comme des énoncés définitifs sur les sujets traités, mais plutôt comme des rapports d'étape sur les études en cours.

Les documents de recherche sont publiés dans la langue officielle utilisée dans le manuscrit envoyé au Secrétariat.

Ce document est disponible sur l'Internet à:

http://www.dfo-mpo.gc.ca/csas/

ISSN 1499-3848 (Printed)
© Her Majesty the Queen in Right of Canada, 2003
© Sa majesté la Reine, Chef du Canada, 2003

Canada

Abstract

The retrospective problem involves systematic differences in sequential population analysis (SPA) estimates of stock size or some other quantity in a reference year. The differences occur as successively more data are used for estimation, and appear to be structural biases that result from a mis-specification of the SPA. In some cases the retrospective problem is so severe that the SPA is considered to be too unreliable for stock assessment purposes. This was the case in the 2002 assessment of the fall spawning herring stock in the southern Gulf of St. Lawrence (4T fall herring; see LeBlanc, MacDougall, and Poirier, 2002). There are many possible causes of retrospective patterns, and it is usually difficult in practice to determine which causes are more likely. In this paper we show how to use local influence diagnostics to investigate whether small changes or perturbations to SPA input components such as catches or natural mortalities can remove or reduce retrospective patterns. The plausibility of the perturbations can be used to assess the likelihood that the component is the source of the retrospective pattern. We apply these local influence diagnostics to the 4T fall herring SPA. We show that reasonable changes in the SPA assumption about the relationship between catch per unit effort (CPUE) and stock size is a plausible source of the retrospective pattern. Catches, natural mortality assumptions, or the weighting of SPA residuals in estimation appear to be less plausible sources of the retrospective patterns.

Résumé

Le profil rétrospectif implique des différences systématiques dans les évaluations de la taille des stocks ou d'autres paramètres obtenues avec l'analyse séquentielle de population (ASP) pour une année de référence donnée. Les écarts, qui surviennent lorsque davantage de données sont successivement utilisées dans l'évaluation, semblent être des biais structurels résultant d'une erreur de spécification de l'ASP. Dans certains cas, le profil rétrospectif est tellement marqué que l'on considérera l'ASP comme étant trop incertaine pour l'évaluation des stocks. C'est ce qui s'est produit dans l'évaluation 2002 du stock de harengs reproducteurs d'automne du sud du golfe du Saint-Laurent (4T, hareng d'automne; voir LeBlanc, MacDougall et Poirier, 2002). Le profil rétrospectif peut être causé par de nombreux facteurs, mais il est d'ordinaire difficile dans la pratique de déterminer quelles en sont les causes les plus probables. Dans cet article, nous démontrons comment utiliser les diagnostics d'influence locale pour déterminer si des faibles changements ou perturbations aux données d'entrée pour l'ASP, tels que les prises ou les mortalités naturelles, peuvent éliminer ou réduire les profils rétrospectifs. On peut s'appuyer sur la plausibilité des perturbations pour évaluer la probabilité qu'une donnée d'entrée particulière soit la source du profil rétrospectif. En appliquant le diagnostic d'influence locale à l'ASP des harengs d'automne de 4T, nous démontrons que des changements raisonnables dans les prémisses de l'ASP concernant le rapport entre les prises par unité effort (PUE) et la taille du stock peuvent être la cause du profil rétrospectif. Les prises, les hypothèses relatives à la mortalité naturelle ou la pondération des résidus de l'ASP dans l'évaluation semblent toutefois être des sources moins plausibles de profil rétrospectif.

104-04233/

1. Introduction

The retrospective problem in sequential population analysis (SPA) has received considerable attention in fish stock assessments. SPA is an analytical model of fishery catch data which can be used to estimate stock size. Based on a time series of annual catch numbers-at-age a in year y, denoted as $C_{a,y}$, SPA produces time series estimates of population numbers-at-age, $N_{a,y}$, and other derived quantities such as total biomass and spawner biomass (SSB). Once fishery catch statistics have been compiled in a given year to estimate $C_{a,y}$'s, an SPA can be performed on the updated catch time series to produce new estimates of the $N_{a,y}$'s. The retrospective problem involves a systematic pattern in stock estimates as catches and other stock data are updated. More specifically, let $S_{y,t}$ denote an SPA estimate of stock size in year y based on data up to year $t \geq y$. A retrospective problem is said to exist if successive estimates $S_{y,t}$, $S_{y,t+1}$, $S_{y,t+2}$, and so on deviate systematically in a decreasing or increasing trend. This problem can be so severe that the SPA is considered to be too unreliable for stock assessment.

A substantial retrospective problem exists in the SPA for the fall spawning herring stock in the southern Gulf of St. Lawrence (see Figure 41 in LeBlanc et al., 2003). This stock, found off the north coast of Nova Scotia, Canada, in NAFO Division 4T (see Figure 1), is commonly referred to as 4T fall herring. Details about the SPA for this stock are provided in **Section 3**. Retrospective estimates of total abundance for ages 5-10 (N_+), average fishing mortality for ages 7-10 ($\bar{F}$), and total biomass for ages 5-10 (B_+) are shown in Figure 2 in the **Appendix**. The estimates cover the years $y = 1978, ..., t$, for $t = 1998, ..., 2002$, where t indicates the last year of catches and other fishery data used to estimate population size. The SPA structure we employ closely resembles that of LeBlanc et al. (2003); our retrospective patterns are also similar. Notice that the estimates of N_+ and B_+ for year y usually decrease as t increases and more data are used. For example, the estimate of B_+ in 1998 based on data up to 1998 is 224 KT; if data up to 2002 are used to obtain an updated 1998 estimate, the result is 131 KT, which is almost 50% lower. This consistently decreasing trend indicates a structural bias in the population size estimators caused by model mis-specification (Evans, 1996). The retrospective patterns shown in Figure 2 are severe, but not uniquely so. For example, a haddock stock considered by Sinclair *et al.* (1991) exhibited similar symptoms, while another stock considered in Mohn (1999) possessed a more severe retrospective pattern.

A common perception in stock assessments is that historic estimates of stock size are more accurate than recent ones. This will tend to be true for stocks that are heavily exploited, and for which accurate catch data exists (Pope, 1972). Hence, when a retrospective problem as in Figure 2 is present, the common perception is that current stock size is over-estimated. This is important because current stock size and trends in recent stock size are usually required for fishery management decisions such as setting the total allowable catch (TAC) for next year. If current stock size is over-estimated then this may mean that the TAC will be set too high and not be sustainable. However, as pointed out by Sinclair *et al.* (1991), for certain types of SPA model mis-specifications the historic estimates may be less accurate than the current estimates. Mohn (1999) also demonstrated this using simulated data with model mis-specifications. He showed that certain

types of model mis-specifications tend to compound over time such that adding more data produces more biased estimates of historic stock size. However, even in these situations the current trends in stock size estimated by SPA may be overly optimistic and lead to a TAC that is too high.

Mohn (1999) presented simulation results to explore the types of retrospective patterns that might arise from various SPA model mis-specifications. He also presented *ad hoc* diagnostics to help discriminate between the possible sources of mis-specification that cause the retrospective problem. He applied these diagnostics to a cod stock off the east coast of Canada (i.e. eastern Scotian Shelf cod). One type of diagnostic involved examining the magnitude of perturbations to model components required to remove the retrospective pattern. Mohn (1999) considered simple perturbations that involved adding a common "effect" to part of a model component. For example, he considered catch perturbations of the form

$$C_{a,y}(\omega) = \begin{cases} C_{a,y}, & y < y_o, \\ C_{a,y}\omega\phi_a, & y \geq y_o, \end{cases}$$

where $C_{a,y}$ and $C_{a,y}(\omega)$ are the observed and perturbed catch at age a in year y. The perturbation $\omega\phi_a$ was applied only after some specified year y_o. The magnitude of the perturbation was controlled by ω, while ϕ_a was an age effect that was fixed in all perturbations. This analysis was used to explore whether unreported discarding of catches starting in year y_o could be the source of the retrospective problem. The magnitude and timing of the perturbation required to remove the pattern was used to assess whether discarding was a plausible causal mechanism. Mohn (1999) considered perturbations to other model components as well. The "parameters" of his perturbation analyses were ω and y_o, which he profiled over to find values that removed the retrospective problem. He measured the retrospective problem using

$$\rho = \sum_{y=y_o}^{Y} \frac{S_{y,y} - S_{y,Y}}{S_{y,y}}, \tag{1}$$

where y_o and Y are the first and last years for which the retrospective pattern was assessed. Recall that $S_{y,Y}$ is the estimate of stock size in year y based on all catch and other stock data up to year Y. If the retrospective estimates for year y based on data only up to year y (i.e. $S_{y,y}$) fluctuate randomly about $S_{y,Y}$ then ρ is approximately zero.

For 4T fall herring we assessed the retrospective pattern for $y_o = 1998$ and $Y = 2002$. The values of ρ for N_+, $\bar{F}$, and B_+ are shown in the upper left hand corners of each panel in Figure 2. The ρ's for N_+ and B_+ are nearly identical, while ρ for $\bar{F}$ is larger in magnitude and negative. If ρ is positive for stock size then we expect ρ to be negative for fishing mortality because this is essentially catch divided by population size, and overestimation of population size leads to underestimation of fishing mortality.

In this paper we also use perturbation analyses to diagnose more likely causes of the retrospective pattern. We improve upon Mohn (1999) by using less constrained perturbations that may be more informative; for example, we investigate catch perturbations of

the form

$$C_{a,y}(\omega) = C_{a,y} \times \omega_{a,y}. \tag{2}$$

In (2) we perturb each $C_{a,y}$ separately, whereas Mohn (1999) considered more simple perturbations. We also find perturbations that remove the retrospective pattern. The advantage of searching over a higher dimensional perturbation space is the potential of finding smaller and more realistic perturbations to remove the retrospective pattern than those presented in Mohn (1999).

We use the local influence approach for perturbation analyses. This method is briefly described in **Section 2**, and more fully discussed in Cadigan and Farrell (2002). Similar to Cook (1986), we use basic concepts in differential geometry to study the effect of a perturbation on ρ. Our approach is based on the slope of the perturbation (influence) surface near the origin, which is why it is referred to as local influence. This approach is computationally more convenient for perturbation analyses because it is based only on slopes at the origin which can be computed using the unperturbed SPA parameter estimates. It does not require re-estimation of the SPA at many points in the perturbation space. For example, if we examined the local effects of catch on SPA output for five perturbations like those in (2), e.g. $\omega = 0.9, 0.95, 1, 1.05, 1.1$, and if there were 100 $C_{a,y}$'s for all ages and years to perturb, we would then have to re-estimate the SPA a total of $4^{100} \simeq 1.6 \times 10^{60}$ times.. This is practically impossible in most situations. The local influence approach we use is much more feasible, but our diagnostics are useful only when the perturbation surface of ρ around a relevant neighborhood of the origin is fairly linear. When the influence surface has substantial nonlinearity then second-order properties such as the local curvature can be investigated, although we do not pursue this here. We show in **Section 4** that for some perturbations of the 4T herring SPA the influence surface of ρ appears to be reasonably linear. We have also observed this in analyses of the cod stock considered in Mohn (1999). We find the direction of maximum slope for ρ at the perturbation origin which is also the local direction of greatest change in ρ. We then examine perturbations in this direction that reduce ρ to near zero. This involves re-estimating the SPA a small number of times, but the directions themselves only require the unperturbed SPA parameter estimates. When the perturbation surface of ρ is linear, our method finds the smallest perturbation that removes the retrospective pattern, as measured by ρ. In **Section 4**, we consider this for four distinct perturbation schemes on catches, mortality, survey catchability, and case weights in order to determine if the retrospective pattern in the 4T fall herring SPA is more likely caused by any of these components.

2. Local Influence Approach

Cadigan and Farrell (2002), hereafter referred to as C&F, considered local influence diagnostics for problems that involved estimating a $p \times 1$ parameter vector θ by maximizing a fit function $l(\theta)$ that had basic smoothness properties. The estimate of θ, denoted as $\hat{\theta}$, was the solution to

$$\dot{l}(\hat{\theta}) = \left.\frac{\partial l(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}} = 0.$$

They perturbed model components with a $k \times 1$ perturbation vector ω and studied the influence of the perturbations on key model results. The perturbation ω was of the form $\omega = \omega(h) = \omega_o + hd$, where ω_o was the null perturbation, d was a fixed direction vector of length one, and h was a scalar that determined the magnitude of the perturbation. Usually for multiplicative perturbations $\omega_o = 1$, and for additive perturbations $\omega_o = 0$. The dimension of k could be large; for example, in one of the SPA applications in C&F, k was equal to 520.

C&F measured influence on $g(\hat{\theta})$, an important but arbitrarily specified scalar SPA result. They considered some basic geometric properties of the influence surface of $g_\omega(\hat{\theta}_\omega)$ versus ω near the origin, ω_o. The primary diagnostic used by C&F was the local slope in the direction d,

$$S(d) = \left.\frac{\partial g_\omega(\hat{\theta}_\omega)}{\partial h}\right|_{h=0} = d' \left.\frac{\partial g_\omega(\hat{\theta}_\omega)}{\partial \omega}\right|_{\omega=\omega_o} = d'\dot{g}_o.$$

A particularly interesting diagnostic was the direction of maximum slope,

$$s_{\max} = \dot{g}_o / \sqrt{\dot{g}_o' \dot{g}_o}. \tag{3}$$

Further computational details are given in **Section 2.1** of C&F.

Some minor modifications of the methods in C&F are required to study local influence for ρ. The influence measure $g_\omega(\hat{\theta}_\omega)$ used by C&F involved parameter estimates from the optimization of a single fit function, whereas ρ is based on parameter estimates from multiple optimizations. For example, the ρ's in Figure 2 are based on five optimizations. Let $S_{y,t}(\omega)$ denote a perturbed estimate of stock size. Using $S_{y,t}(\omega)$'s in (1) gives the perturbed value, $\rho(\omega)$. Note that $S_{y,t}(\omega)$ is a particular form for $g(\hat{\theta})$ using the notation in C&F. Let

$$\dot{S}_{oy,t} = d' \left.\frac{\partial S_{y,t}(\omega)}{\partial \omega}\right|_{\omega=\omega_o}$$

be a vector of local derivatives. Each $\dot{S}_{oy,t}$ can be computed using the methods in C&F. The local slope of ρ_ω in the direction d is given by

$$\dot{\rho}_o = d' \left\{ \sum_y \left(S_{oy,Y} S_{oy,y}^{-1} \dot{S}_{oy,y} - \dot{S}_{oy,Y} \right) S_{oy,y}^{-1} \right\}, \tag{4}$$

where Y is the last year under consideration. Let $\dot{\rho}_{\max}$ denote the maximum local slope. The direction of maximum slope, $s_{\max}$, can be found using (3) with $\dot{g}_o$ equal to the $\{\cdot\}$ term in (4). Additional computational information is provided in the **Appendix** (see **Section 6**).

3. SPA for 4T fall herring

We used essentially the same SPA formulation for 4T fall herring as outlined in LeBlanc et al. (2003). Specifically, we employed the cohort model

$$N_{a,y} = \left(N_{a+1,y+1} e^{M/2} + C_{a,y} \right) e^{M/2}, \tag{5}$$

where $M = 0.2$ is the annual mortality rate due to sources other than the reported catch. The catches are presented in Table 26 in LeBlanc et al. (2003). Equation (5) can be used to estimate the $N_{a,y}$'s for all ages and years if the $N_{A,y}$'s (numbers at the oldest age A for all years) and the survivors in the last year Y, the $N_{a,Y}$'s, are both known. The 4T fall herring SPA covered the years 1978 to 2002, and ages 4 to 10; that is, $Y = 2002$ and $A = 10$. Similar to LeBlanc et al. (2003), all of the $N_{A,y}$'s were approximated using assumptions about fishing mortalities; that is, $F_{10y} = (F_{8y} + F_{9y})/2$, where

$$F_{a,y} = \log(\frac{N_{a,y}}{N_{a+1,y+1}}) - M.$$

Hence, the only cohort parameters to estimate were $N_{4,2002}, ..., N_{9,2002}$.

A catch per unit of effort (CPUE) index of stock abundance was used to estimate survivors (see LeBlanc *et al.*, 2003, Table 33). The CPUE-at-age (i.e. $R_{a,y}$) was assumed to be proportional to stock abundance; that is, $R_{a,y} \approx q_a N_{a,y}$, where q_a is the unknown age-dependent catchability coefficient of the index. The catchabilities for ages 4 to 10 were also estimated. The total number of parameters (θ) to be estimated was 13 for the seven catchabilities and six $N_{a,Y}$'s. The fit function was

$$l(\theta) = \sum_{a,y} \left[\log(R_{ay}) - \log(q_a) - \log\left\{N_{ay}(t)\right\}\right]^2, \tag{6}$$

where $\log\left\{N_{a,y}(t)\right\} = \frac{1}{4}\log(N_{a,y}) + \frac{3}{4}\log(N_{a+1,y+1})$ was a cohort approximation of log September abundance, which coincides with the timing of the fishery used to derive the CPUE index.

LeBlanc et al. (2003) extended their SPA beyond age 10 to include some plus group catches; hence, some of their stock size estimates cover ages greater than 10. We did not include this step; hence, estimates in Figure 2 only cover ages 5-10. Note that estimates at age 4 in the last year are more imprecise compared to other ages; therefore, it is common to exclude this age when examining retrospective patterns. Estimates of survivors are not affected by including the plus group catches in the manner of LeBlanc et al. (2003). Thus, it seems unlikely that including these catches would significantly alter the results of the influence analyses, or lead to different conclusions.

4. Results

4.1. SPA for 4T fall herring

We present some summaries of the SPA estimates of 4T fall herring stock size in Figure 3 based on the entire time series of catch and survey data. Estimates of recruitment ($N_{5,y}$) and total abundance ($\sum_{a=4}^{10} N_{a,y}$) for the entire time series are plotted in the top panel. Estimates of total biomass ($\sum_{a=4}^{10} w_{a,y}N_{a,y}$) are plotted in the bottom panel. Note that $w_{a,y}$ is the beginning-of-year weight-at-age. Some trends in average F's for ages 7-10 and ages 4-6 are shown in Figure 4. The estimated CPUE catchabilities ($\times 10^3$) are shown in Figure 5. The q's increase until ages 7 or 8, then decline slightly.

Standardized residuals based on (6) using the estimated q's and N's are shown in Figure 6. The residuals are standardized by their estimated standard deviation. In the

top panel, residual trends are apparent after 1995. Age dependent patterns (middle panel) in the residuals are not apparent. There appears to be a decrease in the residual variation as $\hat{R}_{ay}$ increases (bottom panel), although the trend is not substantial. More detailed time series plots of the residuals are shown in Figure 7. Age dependent time trends are apparent. Age and year patterns in residuals are also evident in Figure 8. These residual plots suggest that assumptions are mis-specified in the 4T fall herring SPA. This is also evident in Figure 2.

Retrospective patterns in abundance at age estimates ($\hat{N}_{ay}$) are shown in Figure 9. The trends are similar at all ages except age four, where the retrospective patterns are less pronounced. Retrospective patterns in the model predicted total CPUE indices ($R_y = \sum_a R_{ay}$) are shown in Figure 10. In 1998 the SPA did not fit R_{1997} and R_{1998} very well. In 2001 and 2002 the SPA could not match the decline in R_y and this has lead to an increasingly worse fit after 1995.

We have shown residual and retrospective plots that suggest the 4T fall herring SPA is mis-specified; however, these graphics provide little insight about the source of the model mis-specification. In the next four sections we present diagnostics that can assist in identifying the source of the mis-specification.

4.2. Catch perturbations

In this section we present the direction of maximum slope for ρ, i.e. $s_{\max}$, based on catch perturbations described by equation (2). We do this using ρ's for the retrospective patterns in N_+, $\bar{F}$, and B_+. We also use the $s_{\max}$ perturbation of B_+ to find a catch time series that results in reduced retrospective patterns. Let $\omega(h) = \omega_o + hs_{\max}$ and let $\rho_\omega(h) = \rho\{\omega(h)\}$ denote the ρ value based on $\omega(h)$ perturbed catches, $C_{a,y}(h) = C_{a,y} \times \{1 + hs_{a,y}\}$ where $s_{a,y}$ is the element of the vector $s_{\max}$ that corresponds to $C_{a,y}$. Note that the influence graph of $\rho(\omega)$ is a $k+1$ dimensional surface where $k = \dim(\omega)$, and $\rho_\omega(h)$ is called the lifted line that is the intersection of $\rho(\omega)$ and the vertical plane at $s_{\max}$ (or the plane containing all vectors orthogonal to $s_{\max}$). If the influence surface in the direction of $s_{\max}$ is linear, then $\rho_\omega(h) = \rho + \dot{\rho}_{\max} h$, which implies that $\rho_\omega(h) = 0$ when $h = h_{\max} = -\rho/\dot{\rho}_{\max}$. In this case using $\omega(h_{\max})$ gives the smallest perturbation to catches that remove the retrospective pattern, as measured by ρ. When the influence graph has some nonlinearity, $\omega(h_{\max})$ may not completely remove the retrospective patterns. In this situation, a value close to $h_{\max}$ will usually be a better choice for h instead of $h_{\max}$.

The elements of the three $s_{\max}$'s are shown in Figure 11. They are similar for N_+ (panel a) and B_+ (panel c). At the top of each panel we show $-h_{\max}^{-1} = \dot{\rho}_{\max}/\rho$ in percent. Values for $\dot{\rho}_{\max}$ can be derived from $h_{\max}$ and the respective ρ's shown in Figure 2. By construction the $\dot{\rho}_{\max}$'s are always positive, which indicates that perturbations in these directions will increase ρ. Hence, these results suggest that increasing reported catches from the 1993-1996 cohorts, increasing catches at age 10 prior to 2001, and decreasing most other catches will result in relatively large reductions in retrospective patterns. The elements of $s_{\max}$ for $\bar{F}$ tend to be opposite in sign to those for N_+ and B_+, although not always. This is because the retrospective trends in $\bar{F}$ ($\rho = -3.36$) are the reverse of the trends in N_+ ($\rho = 1.27$) and B_+ ($\rho = 1.30$). Changes to the catches in the direction of

$s_{\max}$ in panel b of Figure 11 will increase ρ towards zero and reduce the retrospective pattern in $\bar{F}$.

The practical utility of $s_{\max}$ for finding perturbations that remove or substantially reduce retrospective patterns depends on the linearity of the influence surface. To investigate this, in Figure 12 we plot the percent change in the three ρ statistics, $100 \times \{\rho(\omega) - \rho\}/\rho$, based on perturbations to some individual catches and perturbations to all catches using $s_{\max}$. We used equation (2) for the perturbations, with $\omega = 1 + h$ for individual catch perturbations and $\{\omega_{a,y}\} = 1 + hs_{\max}$ for perturbations in the direction $s_{\max}$. The value of h controlled the amount of perturbation; for example, $h = -0.5$ with an individual catch perturbation meant that the catch was reduced by 50%, whereas $h = 1$ meant that the catch was doubled. The results suggest that there is some nonlinearity in the influence curves. Perturbations in the direction of $s_{\max}$ deviate from a linear trend when $|h| > 0.5$.

If we reduce catches in the direction of $s_{\max}$ in Figure 11 (panel c) by $h_{\max} = -5.6$ then ρ should be near zero for this perturbation. However, as is evident in Figure 12 the influence surface in the direction $s_{\max}$ is somewhat nonlinear and we found that a perturbation with $h = -5.6$ was infeasible because it resulted in negative catches. The smallest value for h we could use was $h = -2.5$. These perturbations are shown in Figure 13; they are simply the results in panel c of Figure 11 scaled by -2.5. The change in most catches is less than $\pm 25\%$, but a small number of catches are altered by more than $\pm 50\%$. Differences between the total observed and perturbed catches are shown in Figure 14. Perturbed catches are substantially larger in total after 1997. The perturbed retrospective patterns are shown in Figure 15. These catch perturbations reduced the retrospective patterns in N_+ and B_+. The ρ statistics dropped from 1.27 to 0.41 for N_+, and from 1.3 to 0.47 for B_+. However, retrospective patterns in N_+ and B_+ still exist, and the retrospective pattern in $\bar{F}$ was worse; that is, ρ decreased from -3.4 to -6.0. The catch perturbations resulted in a poorer fit; the mean square error (MSE) increased from 0.22 to 0.29. These results suggest that small changes to the catches cannot account for the retrospective problem.

4.3. M perturbations

The M perturbations we considered were of the form

$$M_{\omega\, a,y} = M \times \omega_{a,y},$$

where $\omega_o = 1$. These are multiplicative perturbations of M, which was set at 0.2. This is the same value used in LeBlanc et al. (2003).

The elements of the direction vectors are shown in Figure 16. They are broadly similar to the catch perturbation results in Figure 11 for N_+ and B_+, although the influence of M and C at age 10 differs. The results in Figure 16 suggest that an increase in M for the 1992-1996 cohorts and a decrease in M for most other cohorts will reduce the retrospective pattern. The influence curves based on selected M perturbations appear linear (see Figure 17), although a small departure from linearity is evident in the $s_{\max}$ direction when $|h| > 0.75$. The results for B_+ suggests that a perturbation in the direction

shown in panel c of Figure 16 and h around $h_{\max} = -8.5$ will reduce ρ to near zero. After a few trials we chose $h = -6.5$. The trials involved selecting a value for h near -8.5, changing the values for M using $\{M_{a,y}\} = 0.2 \times (1 + hs_{\max})$, and then re-computing ρ to see if it was close to zero. The value of h we selected was smaller in absolute value than the local approximation of -8.5 because of the nonlinearity in the influence curve. Panel c of Figure 16 shows that ρ decreases faster than the linear approximation prediction based on M perturbations in the direction of $s_{\max}$, which is why $h = -6.5$ resulted in a smaller $|\rho|$ than $h_{\max}$. The perturbations based on $h = -6.5$ are shown in Figure 18. The perturbations are large; that is, many of them are greater than 50% in absolute value, and some perturbations are greater than 150%. The perturbed retrospective patterns are shown in Figure 19. They are substantially reduced for N_+, $\bar{F}$, and B_+ compared to those in Figures 2 and 15, although the relative magnitude of the perturbations required to give these reductions is large (i.e. $|h| = 6.5$).

The MSE for the perturbed SPA dropped slightly from 0.22 to 0.2. This is another difference compared to the catch perturbations in Section 4.2. The M and catch perturbed SPA's give substantially different estimates of population size; for example, the maximum absolute difference in total annual biomass for ages 4-10 from the catch and M perturbed SPA's is 260 KT in 2002, which is 54% of the unperturbed value.

4.4. Catchability perturbations

Another interesting perturbation scheme involves the survey catchabilities. A common assumption used in the 4T fall herring SPA is that the CPUE indices are proportional to absolute stock numbers-at-age. The constant of proportionality, q, was assumed to depend on age but not year. This assumption may not be true; for example, it is possible that CPUE catchability changes over time due to changes in fishing practices, etc. Mohn (1999) showed that violations of the constant catchability assumption could cause retrospective patterns. To assess the potential for this we examined influence diagnostics for multiplicative q perturbations,

$$q_{\omega\, a,y} = q_a \omega_{a,y}, \tag{7}$$

where $\omega_o = 1$. This involved perturbations to unknown model parameters. We estimated the q_a's in (7) using the perturbed fit function,

$$\begin{aligned} l_\omega(\theta) &= \sum_{a,y} \left[\log(R_{ay}) - \log(q_{\omega\, a,y}) - \log\left\{N_{ay}(t)\right\}\right]^2 \\ &= \sum_{a,y} \left[\log(R_{ay}) - \log(\omega_{a,y}) - \log(q_a) - \log\left\{N_{ay}(t)\right\}\right]^2. \end{aligned}$$

The perturbed catchabilities were estimated as $\hat{q}_{\omega\, a,y} = \hat{q}_a \omega_{a,y}$. However, the $\omega_{a,y}$'s were fixed and obtained from the local influence diagnostics.

The elements of $s_{\max}$ for the three retrospective measures (ρ's) are shown in Figure 20. They suggest that a trend in the CPUE catchability in 1996-2002 can reduce the retrospective pattern. Global influence curves are shown in Figure 21. As with the M perturbations, the influence curve in the direction of $s_{\max}$ is reasonably linear within the

range of h we considered. We applied the $s_{\max}$ perturbation of B_+ with $h = -2$ (see Figure 22). On a relative scale these are much smaller perturbations than those in Figure 18. The estimates of q_a in (7) were almost identical to the unperturbed estimates; for example, the unperturbed $\hat{q}_{10}$ in Figure 5 was 2.562×10^{-3}, while the perturbed estimate was 2.517×10^{-3}. Differences for other ages were of similar or smaller magnitude. The perturbed retrospective patterns are presented in Figure 23. The q perturbations based on $s_{\max}$ for B_+ greatly reduced the retrospective patterns in N_+ and B_+. The q perturbations improved the fit as well, decreasing the MSE from 0.22 to 0.20.

Age specific retrospective patterns based on the q-perturbations are shown in Figure 24. They are substantially improved compared to the unperturbed results shown in Figure 9 for ages 6 and 7 and marginally improved for ages 8-10. For age 5 the retrospective patterns are about the same magnitude but reversed in sign. The q-perturbed retrospective patterns are worse for age 4.

4.5. Case weight perturbations

Many methods for assessing influence involve the perturbation of case weights. A case refers to a term in the sum in (6), and is the SPA squared residual. A case weight is an extrinsic weight of the squared residual in the fit function, which we perturb as

$$l_\omega(\theta) = \sum_{a,y} \omega_{a,y} \left[\log(R_{a,y}) - \log(q_a) - \log\left\{N_{a,y}(t)\right\}\right]^2.$$

This perturbation scheme can be used to assess, for example, the impact of deleting CPUE indices for a particular year. In this section we assess whether changes in case weights can reduce retrospective patterns. The results in Figure 25 suggest that, compared to the catchability perturbations, larger changes to case weights are required to reduce the retrospective problem. This is because the values of $\dot{\rho}_{\max}$ are larger; however, smaller changes to case weights are required compared to C and M perturbations (see Figures 11 and 16). Global influence curves are shown in Figure 26. Figures 27 and 28 demonstrate that case weight perturbations with $h = -5$ are sufficient to remove the retrospective pattern. One of the main effects of the case weight perturbations shown in Figure 27 is to down-weight the 2001 and 2002 CPUE indices of abundance. The importance of this part of the CPUE data on the retrospective pattern is also obvious in Figure 10.

5. Discussion

We have presented a practical methodology based on local influence diagnostics to assess the potential magnitude of changes in SPA inputs required to reduce or remove retrospective patterns. In the context of a data example, we illustrated how these methods can be employed to identify relatively small perturbations to SPA component inputs that result in greatly reduced retrospective patterns. The diagnostics can identify more likely causes of the retrospective patterns, or at least identify some components that are an unlikely source of the patterns. That is, if the smallest perturbation that removes the retrospective pattern is unrealistic then we can reasonably conclude that the pattern is not caused by the component. Although this is obviously not as desirable as identifying

the components that are at the root of the retrospective pattern, the determination of which components that are unlikely to have caused the problem is nevertheless useful.

We have applied the proposed methods to the 4T fall herring SPA, and studied the influence of commercial catches, natural mortality, survey catchability, and estimation case weights on retrospective patterns. We concluded that relatively small perturbations to assumptions about the catchability of CPUE indices could remove the retrospective patterns. Larger changes to case weights or the assumptions about M were required to remove the patterns, while local catch perturbations were not able to completely eliminate them. The plausibility of the perturbations, or the SPA perturbed stock estimates, are best assessed by 4T herring experts who are knowledgeable about the fishery and other scientific information for this stock.

Although our local influence analyses have produced a large amount of output to examine, our main conclusion is based on only the values of $\dot{\rho}_{\max}$ for the four perturbation schemes. For B_+, these slopes in percent of ρ are 18%, 12%, 40%, and 19% for the C, M, q, and case weight perturbations, respectively. These slopes suggest that smaller changes to the CPUE catchability model are required to reduce the retrospective pattern than for the other SPA model components. The N_+ ρ slopes lead to the same conclusion. Note that a different conclusion is reached from the $\bar{F}$ ρ slopes; however, as is evident in Figure 12, there is substantial nonlinearity in the influence surface for this ρ and perturbation scheme, and this reduces the utility of local influence diagnostics.

Before local influence diagnostics are interpreted it is important to check the "near-local" linearity of the influence surface, which we did in Figures 12, 17, 21, and 26. An alternative option is to develop local curvature diagnostics to indicate when the influence surface is locally nonlinear. This is a useful area for future research.

Retrospective-corrected stock scenarios are shown in Figure 29. It is interesting that catch and M perturbations lead to somewhat different estimated stock trajectories, since they both play a very similar role in SPA by accounting for population deaths; however, an important difference in their roles involves the constraints on fishing mortalities at the oldest age A. These constraints are commonly used and often make an SPA more sensitive to errors in catches at age A than similar sized errors in M at that age. It is also interesting that the CPUE q and case weight perturbations resulted in stock trajectories that are very similar to the unperturbed result. The conclusion we draw from this is that a retrospective-corrected SPA may not produce lower estimates of stock size. This is contrary to the common perception that a retrospective problem implies that the SPA will over-estimate current stock size.

We also conducted the retrospective local influence analysis on the data and model structure used in the 2002 assessment of the 4T herring stock (see LeBlanc, MacDougall, and Poirier, 2002). The main difference in the 2002 and 2003 SPA formulations was that in 2002 the CPUE time series was split (1978-1991, 1992-2001) and treated as two different indices. The local influence q-perturbations for ages 5 and greater "attempted" to re-connect the catchabilities for the 1978-1991 and 1992-2001 periods. This suggested that splitting the CPUE index series may have contributed to the retrospective problem. Indeed, when the CPUE time series was not split then the retrospective pattern

was substantially reduced; however, with the addition of the 2002 data the retrospective pattern increased again. Our conclusion from this was that in the "split-q" formulation, some of the retrospective pattern resulted from confounding between stock size and CPUE catchability in the 1992-2001 period. In the "constant-q" formulation (i.e. constant q for 1978-2001) there was less confounding, which in turn reduced the retrospective pattern. However, this did not mean that the "constant-q" formulation was reasonable. All that happened was that the evidence of model mis-specification was transferred from the retrospective plots to the residuals plots; that is, residuals looked worse for the "constant-q" formulation compared to the "split-q" formulation. The 2002 CPUE values were also inconsistent with the SPA, similar to the 2001 values, and this caused the residuals and retrospective patterns to look worse compared to 2002 assessment results; that is, the addition of more inconsistent data to the SPA increased the evidence of model mis-specification. To their credit, LeBlanc et al. (2003) have attempted to conservatively adjust for the mis-specification.

We found that the retrospective diagnostics were fairly similar when ρ was based on different stock quantities (e.g. N_+, $\bar{F}$, or B_+), although we observed some differences in the effect of the perturbations on retrospective patterns. One would expect the diagnostics based on N_+ and B_+ to be similar. Perhaps a better set of diagnostics would involve the numbers of young and old fish instead of N_+ and B_+. In many applications spawner biomass would be an appropriate way to quantify the numbers of old fish. Another factor to consider is the number of years used to measure the retrospective pattern. Also, the retrospective metric proposed by Mohn (1999) and used in this paper (equation 1) is a relative average which may not be appropriate in some cases, especially when some $S_{y,y}$'s are large.

It is possible that retrospective patterns are caused by mis-specifications of two or more model components. For example, the retrospective patterns could be caused by mis-reported catches and incorrect assumptions about CPUE catchabilities. Local influence diagnostics based on perturbations to multiple components are relatively straightforward to implement; however, scaling the perturbations is a problem. For example, multiplicative perturbations to catches may not be comparable to multiplicative perturbations to survey catchabilities. Local influence diagnostics based on multiple component perturbations will likely be very sensitive to the relative scaling of the perturbations to different components. This problem also affects separate perturbations to model components such as the ones we have presented. If the perturbation schemes are not comparable then the local slopes (e.g. $\dot{\rho}_{\max}$) are also not comparable and are subsequently not useful for determining which perturbations are more likely. This is discussed further in C&F.

The local influence diagnostics can be improved by utilizing more realistic perturbation schemes. For example, potential errors in catches may not be entirely multiplicative in nature because the magnitude of errors in large reported landings could be larger or smaller than the errors in small reported landings. Also, completely independent perturbations of M may not be realistic, and smoother perturbations may be more appropriate. Such modifications are relatively straight-forward to implement within the diagnostic framework we have presented.

References

Cadigan, N. G. and Farrell, P. J. 2002. Generalized local influence with applications to fish stock cohort analysis. Applied Statistics, 51: 469-483.

Cook, R. D. 1986. Assessment of local influence (with discussion). Journal of the Royal Statistical Society, Series B, 48: 133-169.

Evans, G. T. 1996. Using the elementary operations of sequential population analysis to display problems in catch or survey data. Canadian Journal of Fisheries and Aquatic Science, 53: 239-243.

Fanning, L. P., Mohn, R. K., and MacEachern, W. J. 1995. An assessment of 4VsW cod in 1994 with consideration of ecological indicators of stock status. DFO Atlantic Fisheries Research Document 95/73. 29 pp.

LeBlanc, C.H., MacDougall, C., and Poirier, G.A. 2002. Assessment of the NAFO 4T southern Gulf of St. Lawrence herring stocks in 2001. DFO CSAS Res. Doc. 2002/053.

LeBlanc, C.H., Poirier, G.A., Chouinard, G., and MacDougall, C. 2003. Assessment of the 4T southern Gulf of St. Lawrence herring stocks in 2002. DFO CSAS Res. Doc. 2003/040.

Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES Journal of Marine Science, 56: 473-488.

Pope, J. G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. ICNAF Research Bulletin, 9: 65-74.

Sinclair, A., Gascon, D., O'Boyle, R., Rivard, D., and Gavaris, S. 1991. Consistency of some Northwest Atlantic ground fish stock assessments. NAFO Scientific Council Studies, 16: 59-77.

6. Appendix

In this section we provide additional information on how we compute local influence diagnostics. We first consider influence for a scalar statistic from an SPA, and then we consider the retrospective ρ's which are scalar statistics computed from more than one SPA. Let $\theta = \{q_4, ..., q_{10}, N_{4,2002}, ..., N_{10,2002}\}$ denote the parameters to estimate in the 4T herring SPA. Also, let $g(\theta) = g(\theta; C, M)$ denote some function of the SPA output, where C and M are age by year matrices of catches and other mortality rates that are inputs to the model. For example, $g(\theta)$ could be the total abundance in year t; that is,

$$g(\theta) = \sum_a N_{a,t}.$$

This $g(\theta)$ is a function of θ and also C and M because $N_{a,t} = N_{a,t}(\theta; C, M)$. Note however that N is not a function of the q's.

Under standard conditions (see **Section 2.1** in C&F) the local slope of a perturbed $g(\theta)$ with respect to the perturbation is

$$\left.\frac{\partial g_\omega(\hat{\theta}_\omega)}{\partial \omega}\right|_{\omega=\omega_o} = \left.\frac{\partial g_\omega(\hat{\theta})}{\partial \omega}\right|_{\omega=\omega_o} + \left.\frac{\partial \hat{\theta}'_\omega}{\partial \omega}\right|_{\omega=\omega_o} \left.\frac{\partial g(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}}, \tag{8}$$

where ω is a vector of all perturbations. This equation is just a simple application of the chain rule for differentiation. We illustrate how to compute this equation using catch perturbations and our total abundance example for $g(\theta)$. Each individual catch is perturbed as $C_{a,y}(\omega_i) = C_{a,y} \times \omega_i$, where $i = 1, ..., AY$ and AY is the total number of ages and years in the SPA. Any convenient mapping between i and a,y can be chosen. Let $\omega = \{\omega_i\}$ where $\dim(\omega) = AY$. The first term on the right hand side (rhs) of (8) is

$$\left.\frac{\partial g_\omega(\hat{\theta})}{\partial \omega}\right|_{\omega=\omega_o} = \sum_a \left.\frac{\partial N_{a,t}(\theta; C_\omega, M)}{\partial \omega}\right|_{\omega=\omega_o}.$$

It is difficult to simplify the N derivative further because of the rather complicated way that catches determine SPA numbers depending on the type of constraints used on fishing mortalities, etc. In our applications we use numerical techniques to compute $\partial g_\omega(\hat{\theta})/\partial\omega$; that is, for the ith element in the above vector of derivatives and for some small δ (e.g. we use $\delta = 10^{-3}$), we compute

$$\left.\frac{\partial g_\omega(\hat{\theta})}{\partial \omega_i}\right|_{\omega_i=\omega_{oi}} \doteq \frac{g_{(-2)} - 8g_{(-1)} + 8g_{(+1)} - g_{(+2)}}{12\delta},$$

where $g_{(\pm z)} = g(\hat{\theta}; C_{\pm z}, M)$ and $C_{\pm z}$ is the observed catch matrix but with the ay-th element equal to $1 \pm z\delta$ times the observed catch. Note that a and y are the age and year that map to i. The derivative approximation is based on a five-point formula and is reasonably accurate (i.e. $O(\delta^4)$). The value for δ we use has yielded satisfactory

results, although better choices for δ may be available in the numerical literature. The unperturbed SPA parameter estimates ($\hat{\theta}$) are used in all the computations. We evaluate $\partial g_\omega(\hat{\theta})/\partial\omega|_{\omega_o}$ using $4 \times AY$ SPA function evaluations.

We use a similar procedure to compute the last term on the rhs of (8), $\partial g(\theta)/\partial\theta|_{\theta=\hat{\theta}}$. This derivative is very fast to evaluate. Note that if the year t that we compute total abundance for is the last year in the SPA then $\partial g(\theta)/\partial\theta|_{\theta=\hat{\theta}} = A$ because $g(\theta)$ is just the sum from 1 to A of some of the θ's. The hard part to compute in (8) is $\partial\hat{\theta}'_\omega/\partial\omega|_{\omega=\omega_o}$.

Fortunately computing $\partial\hat{\theta}'_\omega/\partial\omega|_{\omega=\omega_o}$ does not require computing $\hat{\theta}_\omega$. It can be shown (see C&F) that

$$\left.\frac{\partial\hat{\theta}'_\omega}{\partial\omega}\right|_{\omega=\omega_o} = -\ddot{F}^{-1}\Delta,$$

where F is the SPA fit function used to estimate the θ (i.e. equation 6 for 4T herring), $\ddot{F}$ is the Hessian matrix,

$$\ddot{F} = \left.\frac{\partial^2 F(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\hat{\theta}},$$

and

$$\Delta = \left.\frac{\partial^2 F_\omega(\theta)}{\partial\theta\partial\omega'}\right|_{\theta=\hat{\theta},\omega=\omega_o}.$$

The Hessian will usually be available as a result of the estimation of θ. We also use numerical procedures to compute both $\ddot{F}$ and Δ. If $\dim(\omega)\times\dim(\theta)$ is large (e.g. 500×30) then Δ can take upwards of 30 to 60 minutes to compute with the SPA software that we employ. However, for a simple structure like the 4T herring SPA then Δ can be computed quickly.

The derivative results can be collected to compute $g_o = \partial g_\omega(\hat{\theta})/\partial\omega|_{\omega_o}$. The maximum local slope is $S(s_{\max}) = \sqrt{\dot{g}'_o\dot{g}_o}$, and the direction of maximum slope is given by $\dot{g}_o/S(s_{\max})$ (see **Section 2.1** in C&F).

The local influence analysis of the retrospective ρ statistic is a straight-forward, although somewhat tedious, extension of the methods for influence in a single SPA. One has to define the SPA statistic S that ρ is based on (e.g. see equation 1), and then compute the local slope of the perturbed ρ,

$$\rho(\omega) = \sum_{y=y_o}^{Y} \frac{S_{y,y}(\omega) - S_{y,Y}(\omega)}{S_{y,y}(\omega)},$$

based on the perturbed statistics $S(\omega)$. It is not difficult to show that $\partial\rho(\omega)/\partial\omega|_{\omega=\omega_o}$ is given by (4). Note that the S_o's in this equation are $AY \times 1$ vectors for the catch perturbations. The $\dot{S}_o$'s are obtained using the g_o equations with $g(\theta) = S(\theta)$. The direction of maximum slope is $s_{\max} = \dot{\rho}_o/\sqrt{\sum_i \dot{\rho}_{oi}^2}$, where $\dot{\rho}_{oi}$ is an element of the vector $\dot{\rho}$.

Figures

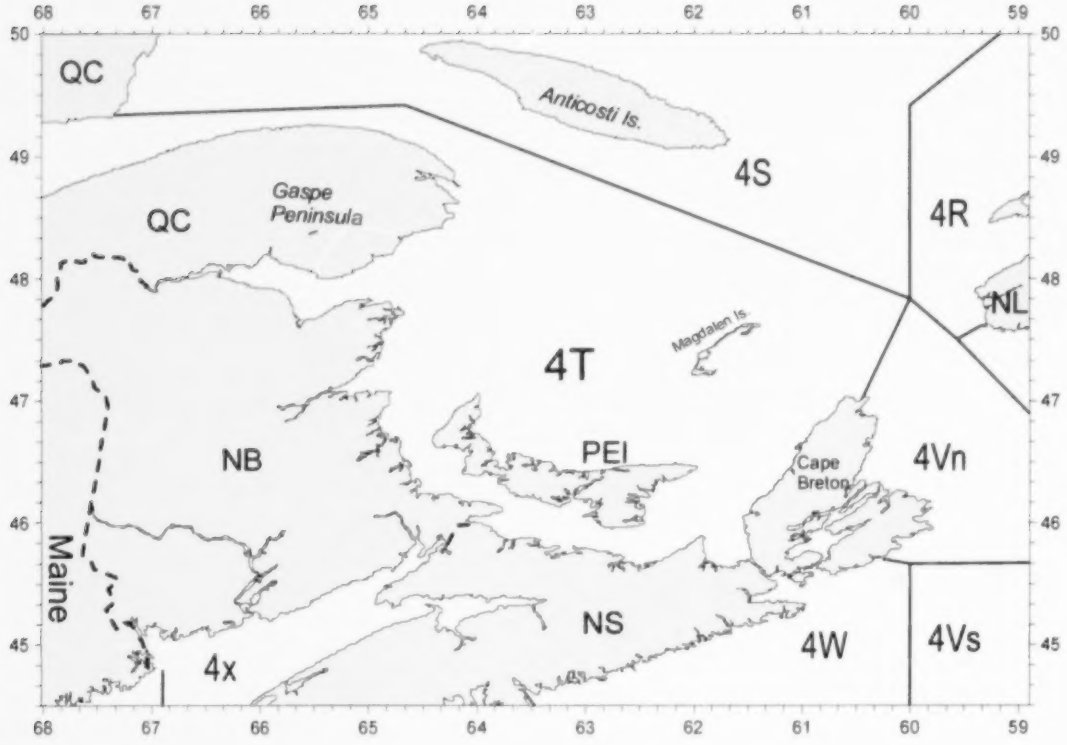

Figure 1: Northwest Atlantic Fisheries Organization (NAFO) Division 4T, located in the southern Gulf of St. Lawrence, off the east coast of Canada.

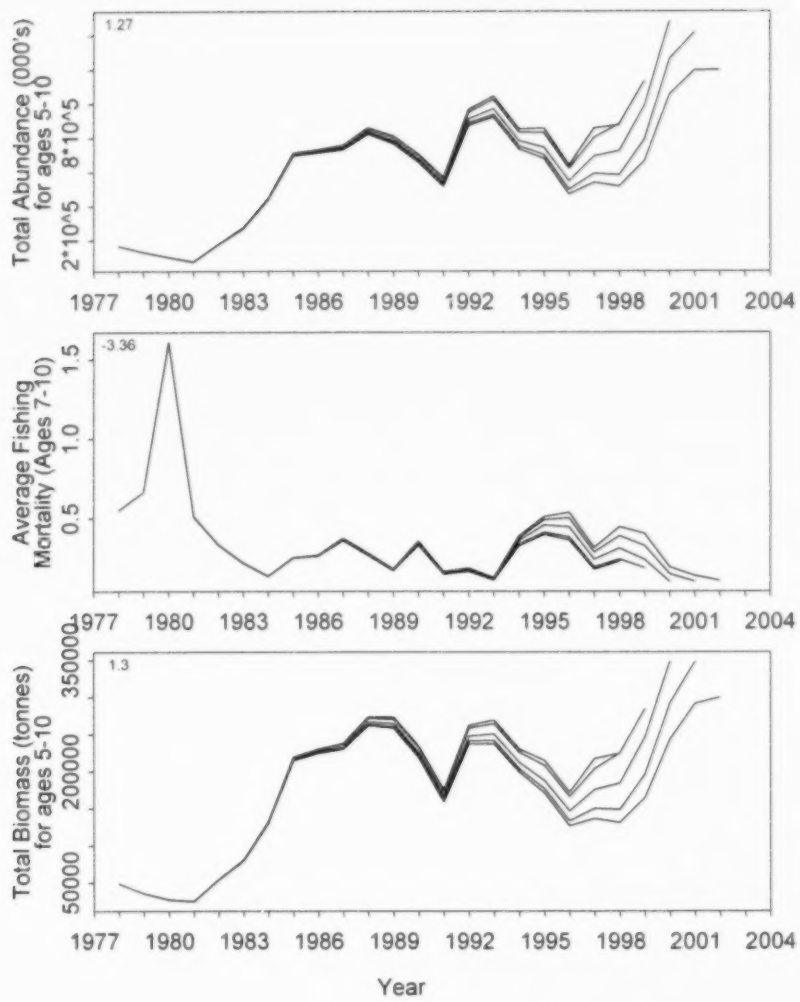

Figure 2: Retrospective estimates for the 4T fall herring stock. The retrospective ρ statistic is shown in the top left-hand corner of each panel. Top panel: N_+, middle panel: $\bar{F}$, bottom panel: B_+.

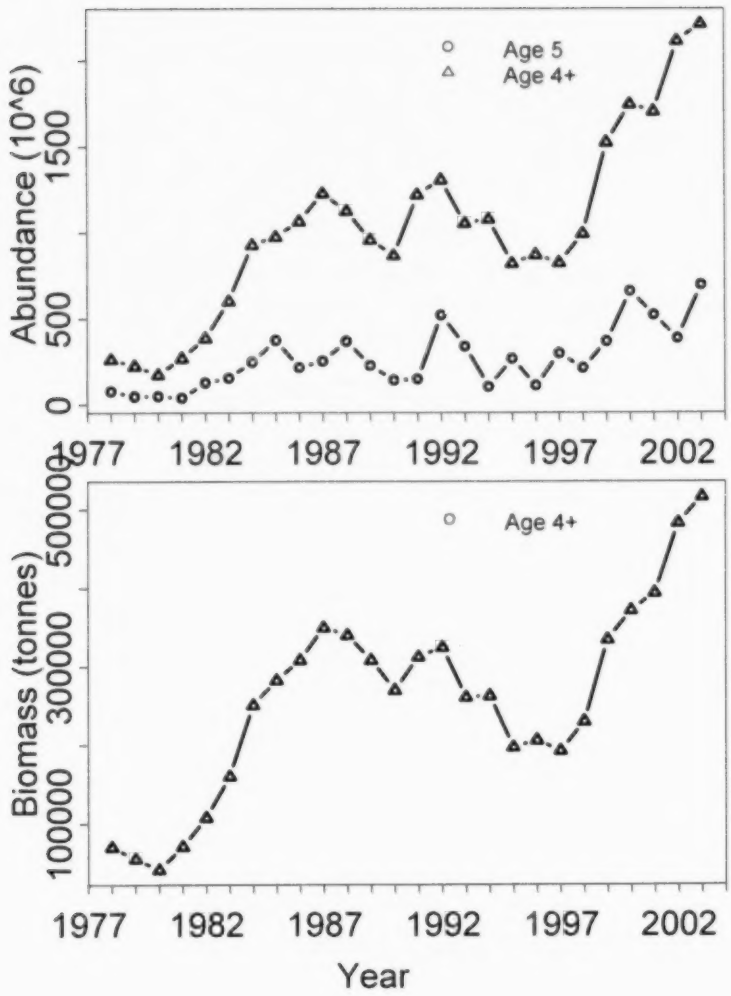

Figure 3: Estimates of recruitment and total abundance at ages 4-10 (top panel), and total biomass at ages 4-10 (bottom panel).

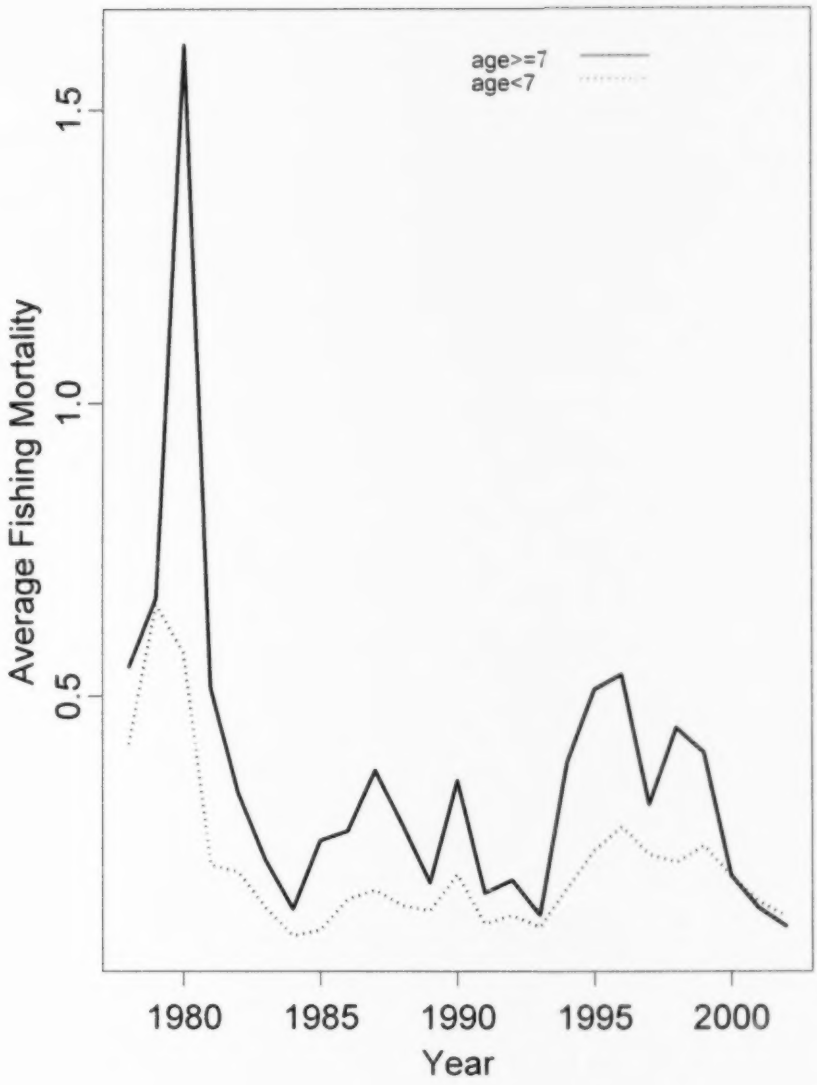

Figure 4: Time series of average estimated fishing mortality for ages 4-6 and ages 7-10.

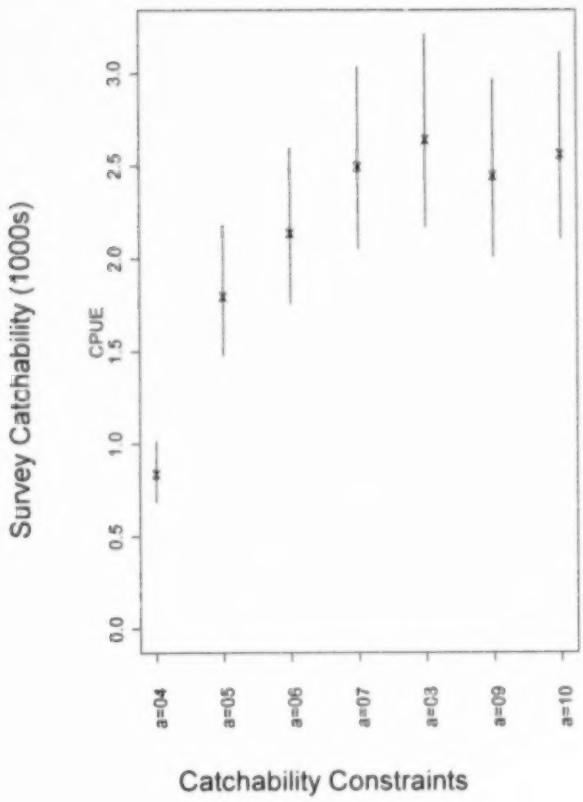

Figure 5: Catch per unit effort (CPUE) SPA catchability estimates (q_a's). Vertical lines represent 95% confidence intervals.

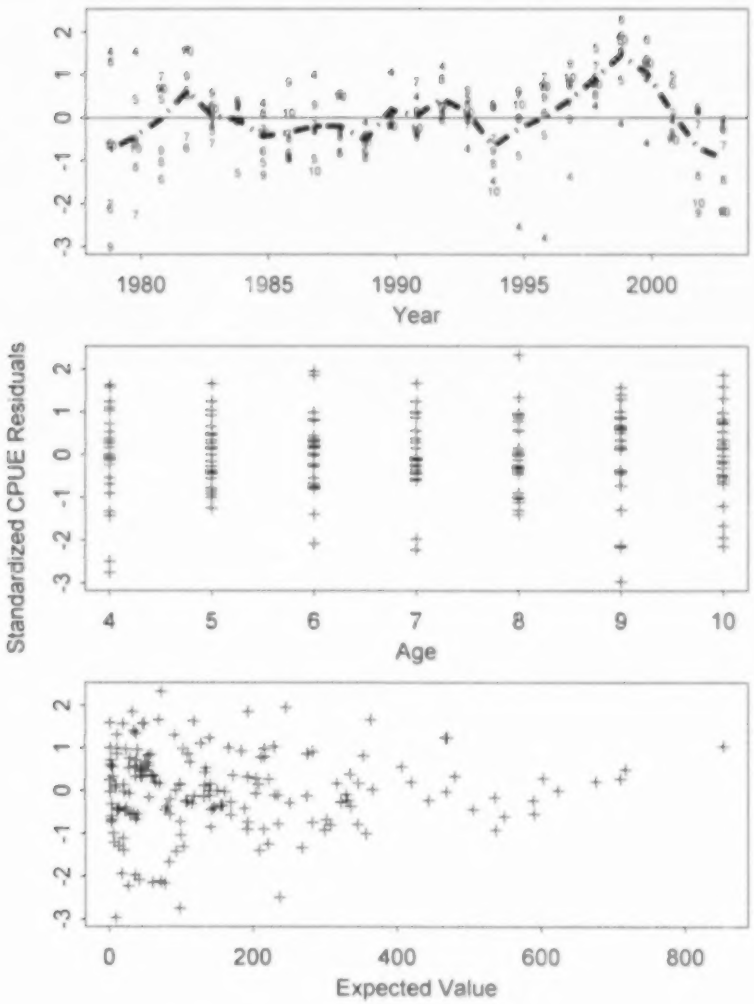

Figure 6: Standardized residuals plotted versus year plus month (e.g. 1999.75, top panel), age (middle panel), and predicted value ($\hat{R}_{ay}$, bottom panel). The dashed line in the top panel connects the average residual each year, and the plotting symbol is the age.

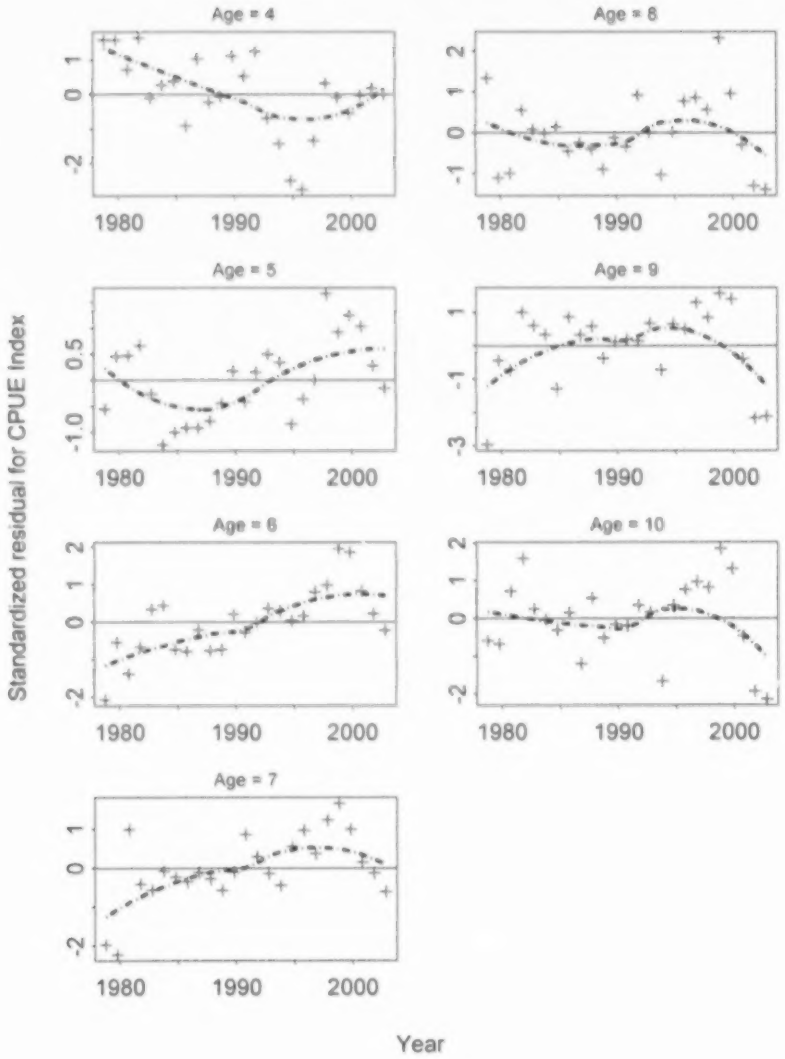

Figure 7: Time series plots of standardized residuals. Time is year plus month (e.g. 1999.75). Each panel shows the CPUE residuals for an age ($a = 4, ..., 10$). The dashed line shows a lowess smoother fit to the residuals.

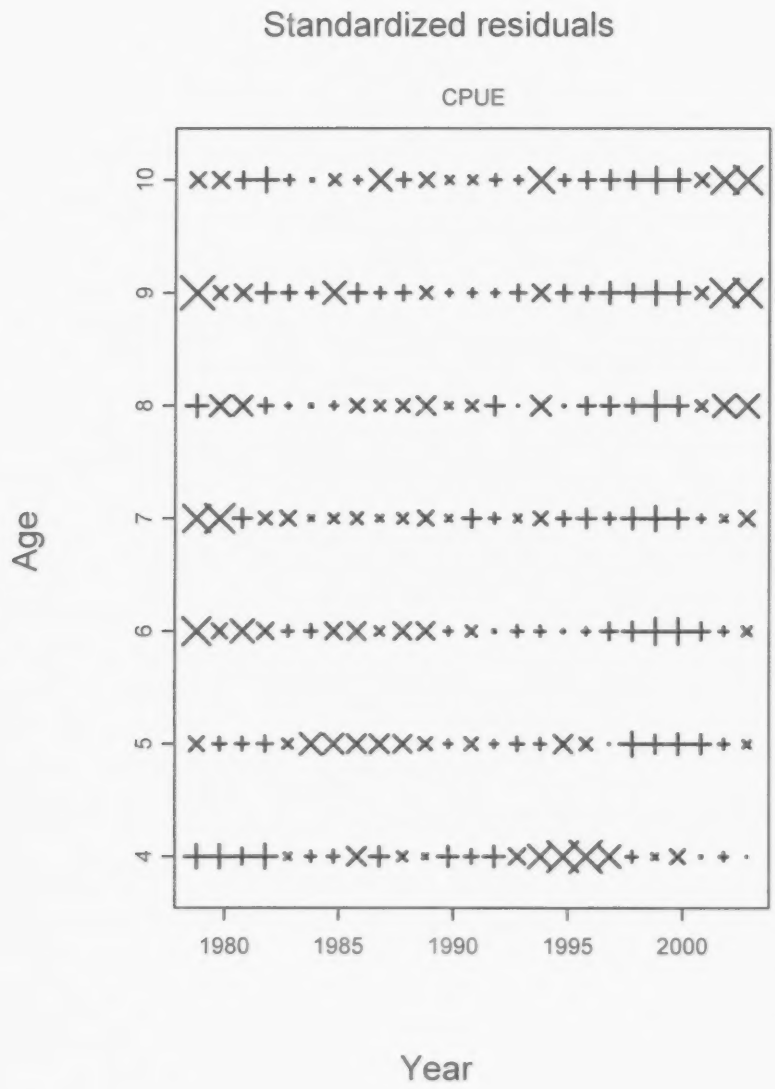

Figure 8: Standardized residuals versus age and year plus month (e.g. 1999.75). The size of the plotting symbols is proportional to the absolute value of the residuals. An × denotes a negative residual, and a + denotes a positive residual.

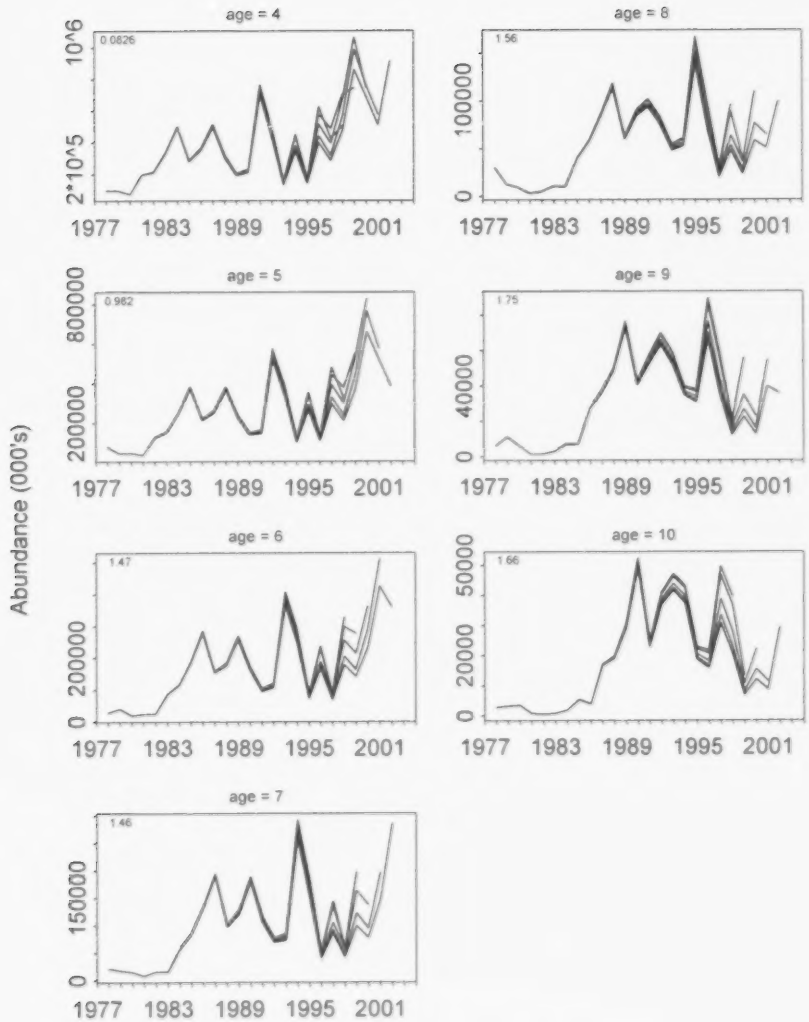

Figure 9: Retrospective plots of 4T herring SPA stock abundance estimates. Each panel shows the results for an age. The retrospective ρ statistic is shown in the top left-hand corner of each panel.

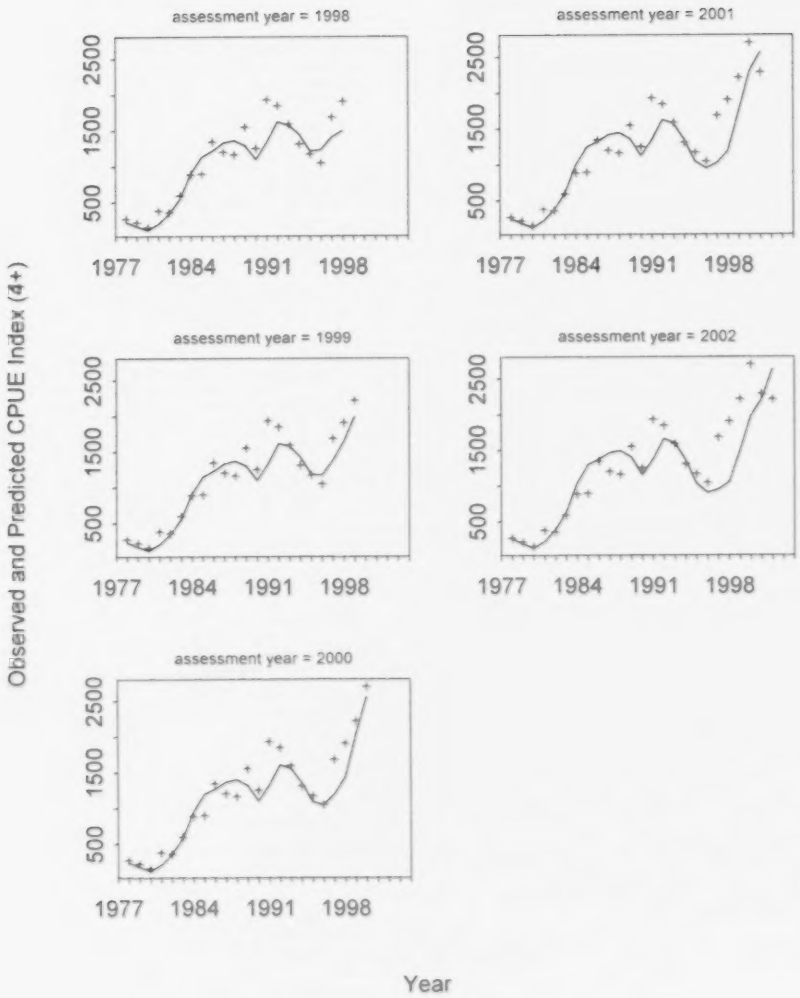

Figure 10: Observed and retrospective predicted total (ages 4-10) CPUE.

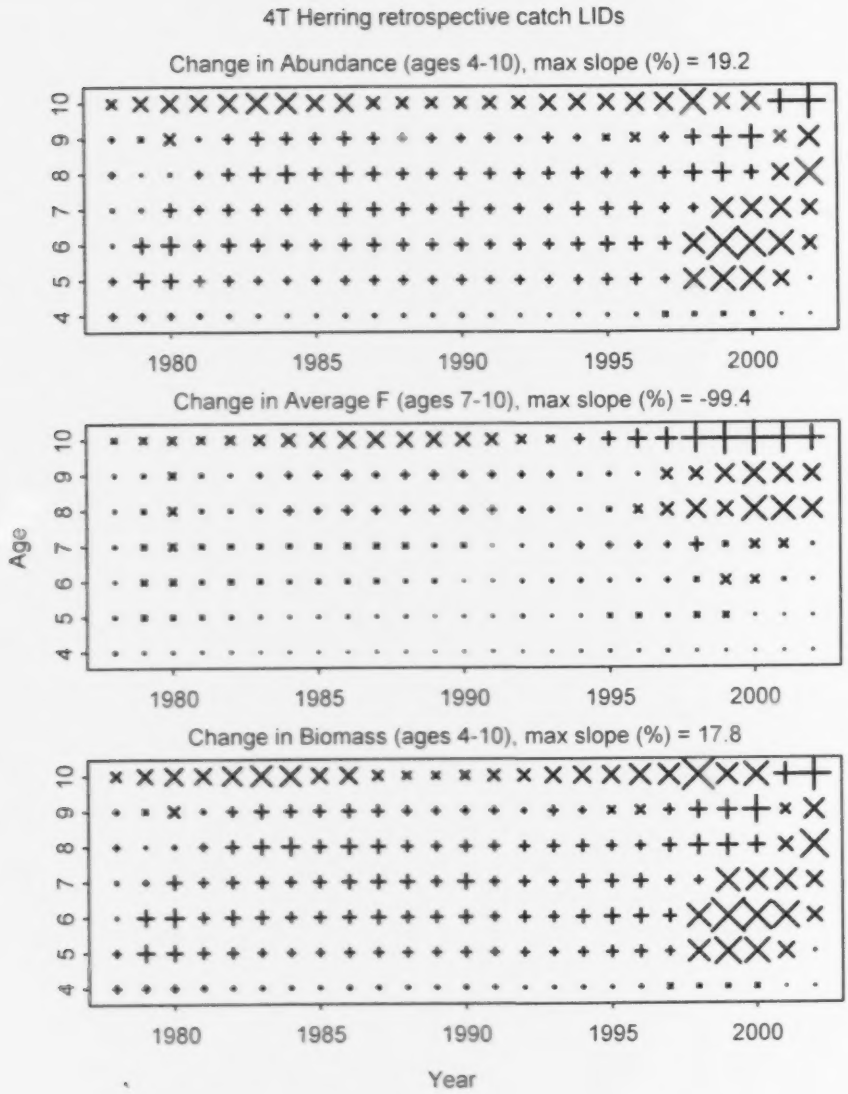

Figure 11: Catch local influence diagnostics (LID's) for ρ. Each panel shows the elements of $s_{\max}$. The size and type of the plotting symbols are proportional to the absolute value and sign of the elements, respectively. Negative is denoted by an $\times$. Panel a: N_+; Panel b: $\bar{F}$; Panel c: B_+. At the top of each panel $\dot{\rho}_{\max}$ is shown in percent of the ρ values in Figure 2.

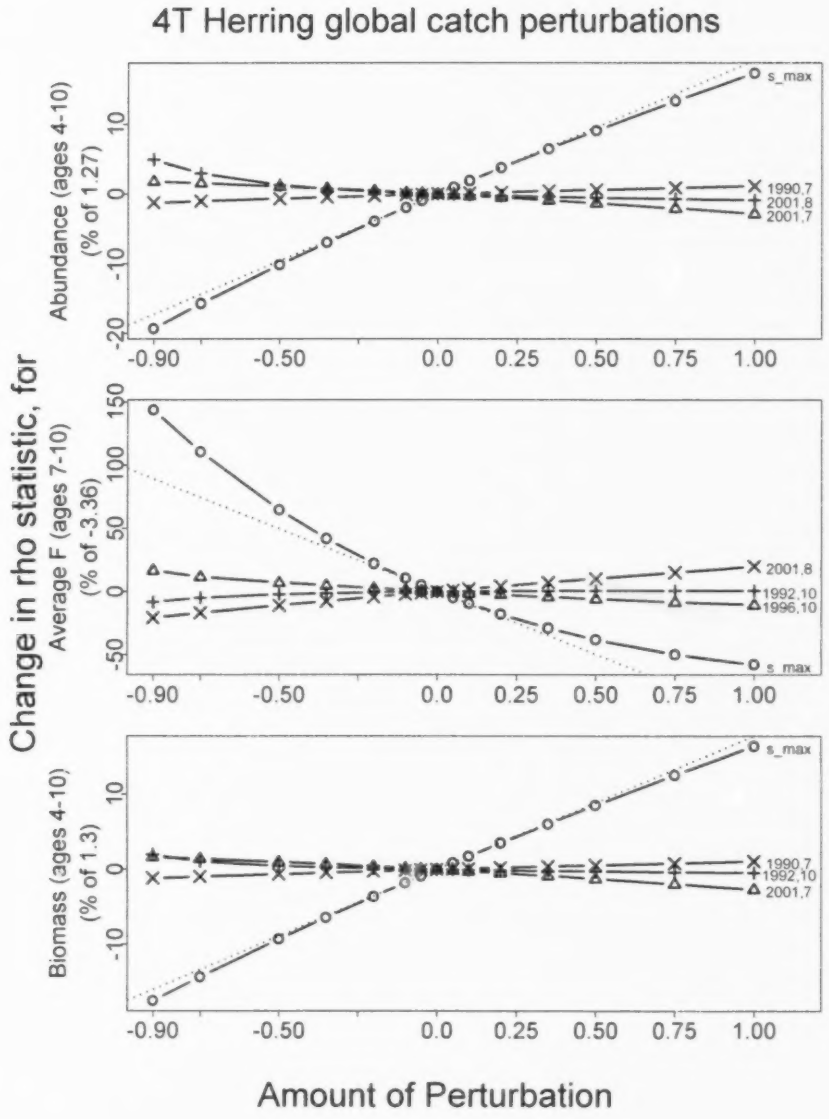

Figure 12: Displacement in ρ's based on some catch global perturbations. All results are in percent of unperturbed estimates. The (year,age) indicates a perturbation of a catch only for that age and year. The dashed line is a straight line with a slope equal to $\dot{\rho}_{\max}$. Panel a: N_+; Panel b: $\bar{F}$; Panel c: B_+.

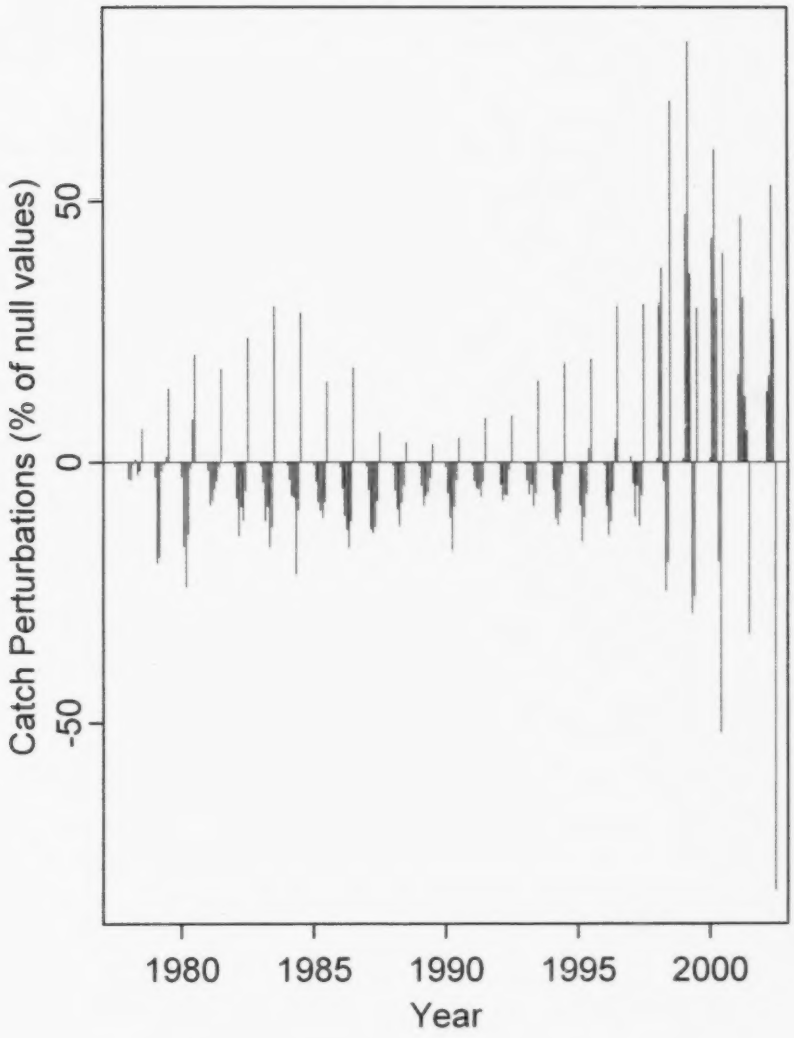

Figure 13: Catch perturbations to reduce the retrospective pattern in B_+. Each vertical line shows the perturbation to a catch. Perturbations are clustered by year and shown sequentially for ages 4-10. The values are $-h \times s_{\max}$ in percent with $h = 2.5$.

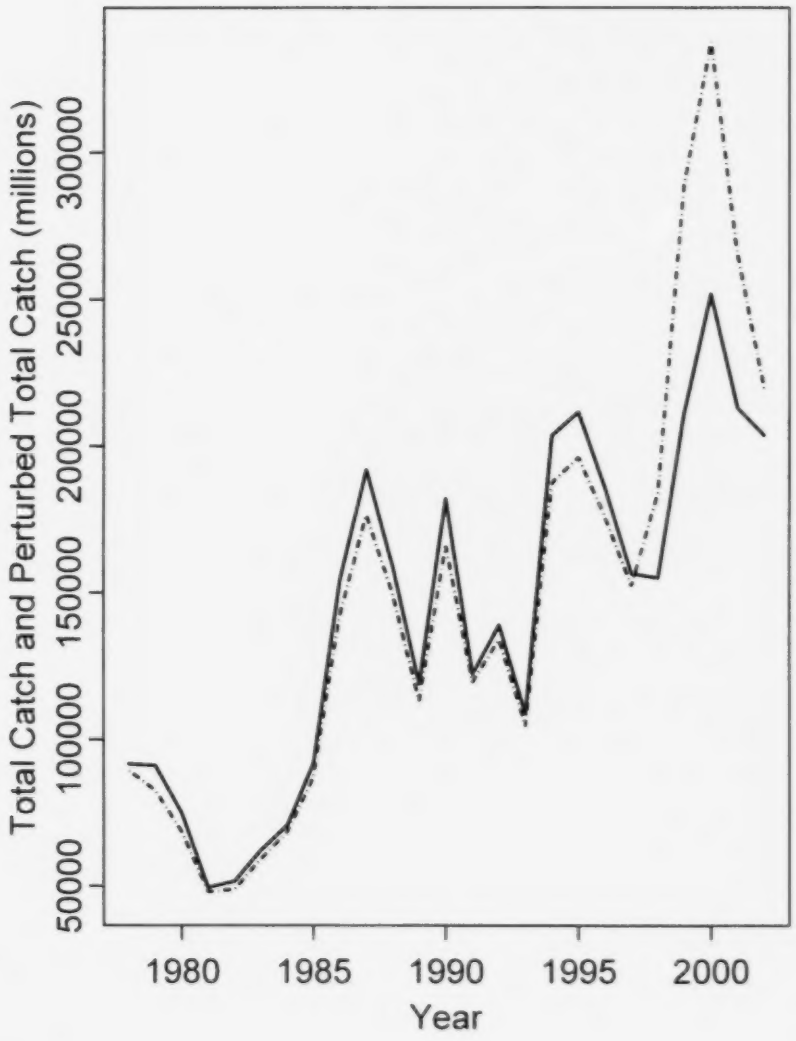

Figure 14: Total observed annual catch (solid line), and total perturbed annual catch (dashed line).

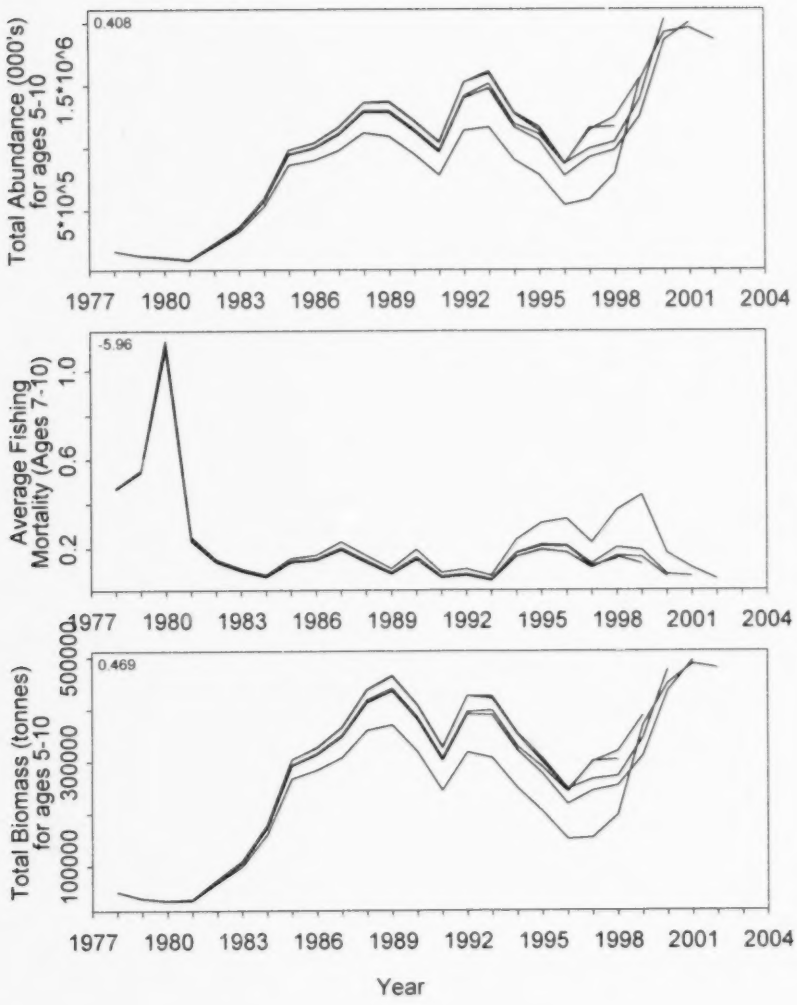

Figure 15: Retrospective estimates based on B_+ $s_{\max}$ perturbed catches with $h = -2.5$. The value of ρ is shown in the top left-hand corner. Top panel: N_+; middle panel: $\bar{F}$; bottom panel: B_+.

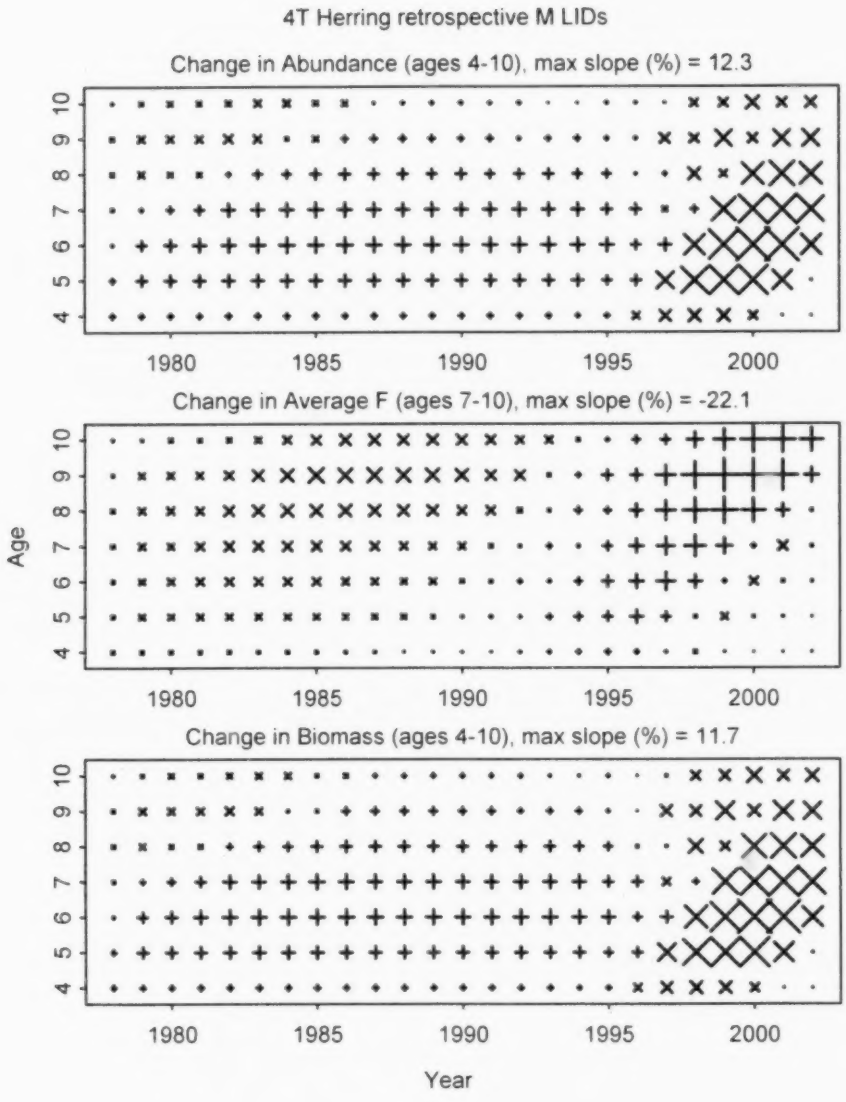

Figure 16: M local influence diagnostics (LID's) for ρ. Each panel shows the elements of $s_{\max}$. The size and type of the plotting symbols are proportional to the absolute value and sign of the elements, respectively. Negative is denoted by an $\times$. Panel a: N_+; Panel b: $\bar{F}$; Panel c: B_+. At the top of each panel $\dot{\rho}_{\max}$ is shown in percent of the ρ values in Figure 2.

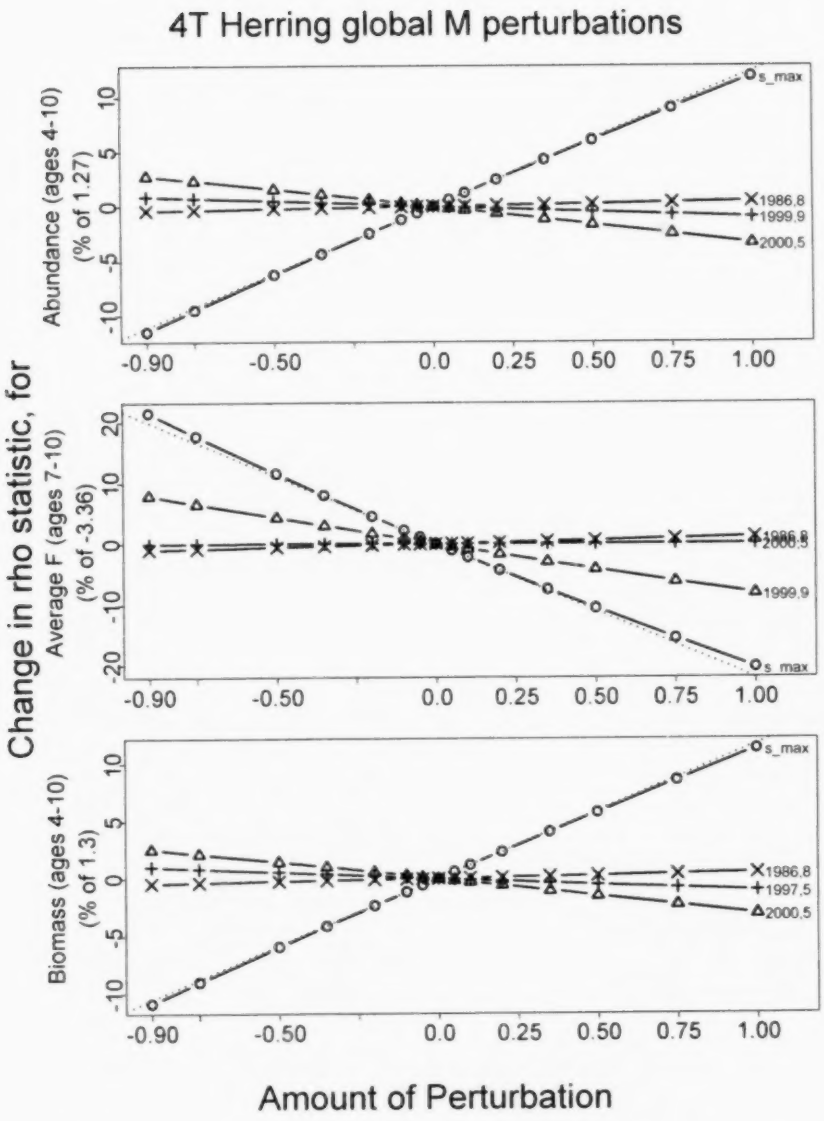

Figure 17: Displacement in ρ's based on some M global perturbations. All results are in percent of unperturbed estimates. The (year,age) indicates a perturbation of M only for that age and year. The dashed line is a straight line with a slope equal to $\dot{\rho}_{\max}$. Panel a: N_+; Panel b: $\bar{F}$; Panel c: B_+.

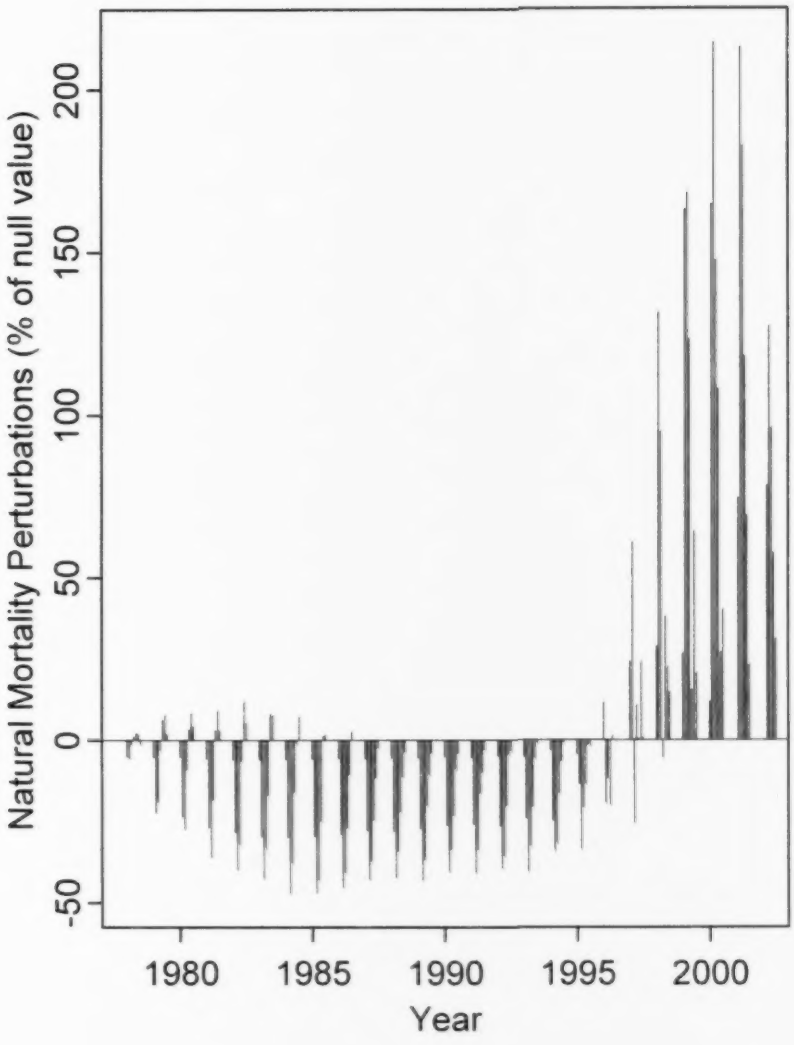

Figure 18: M perturbations to reduce the retrospective pattern in B_+. Each vertical line shows the perturbation to M for that age and year. Perturbations are clustered by year and shown sequentially for ages 4-10. The values are $-h \times d_{\max}$ in percent with $h = 6.5$.

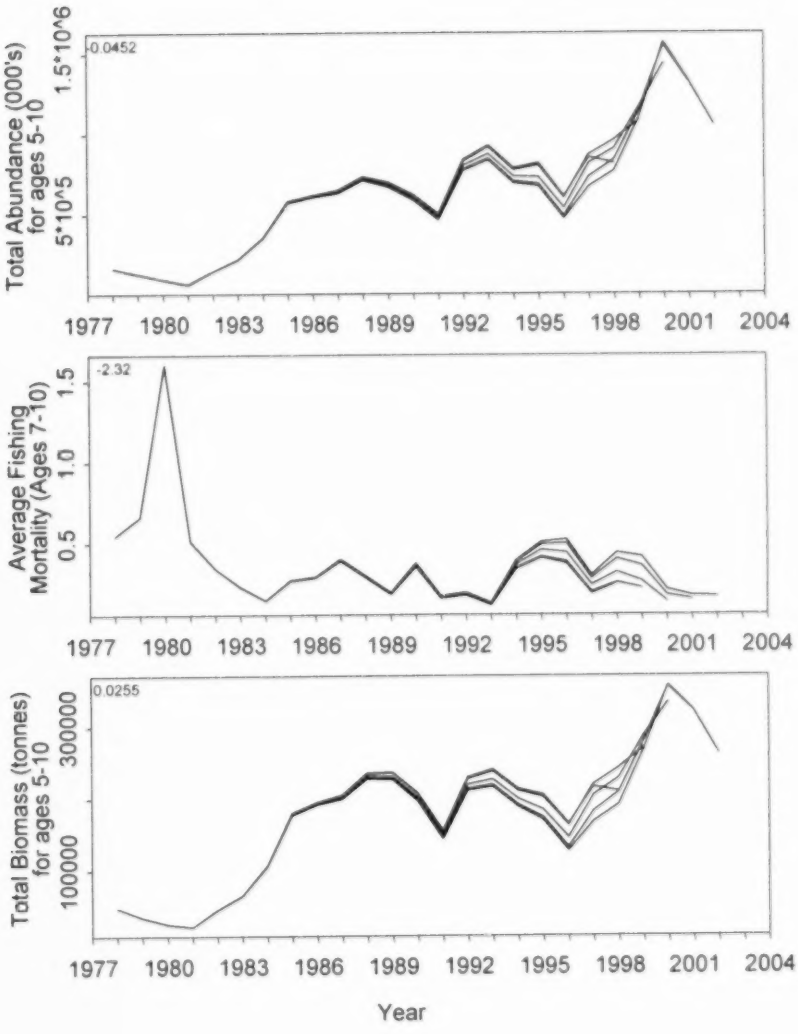

Figure 19: Retrospective estimates based on B_+ $s_{\max}$ perturbed M's with $h = -6.5$. The value of ρ is shown in the top left-hand corner. Top panel: N_+; middle panel: $\bar{F}$; bottom panel: B_+.

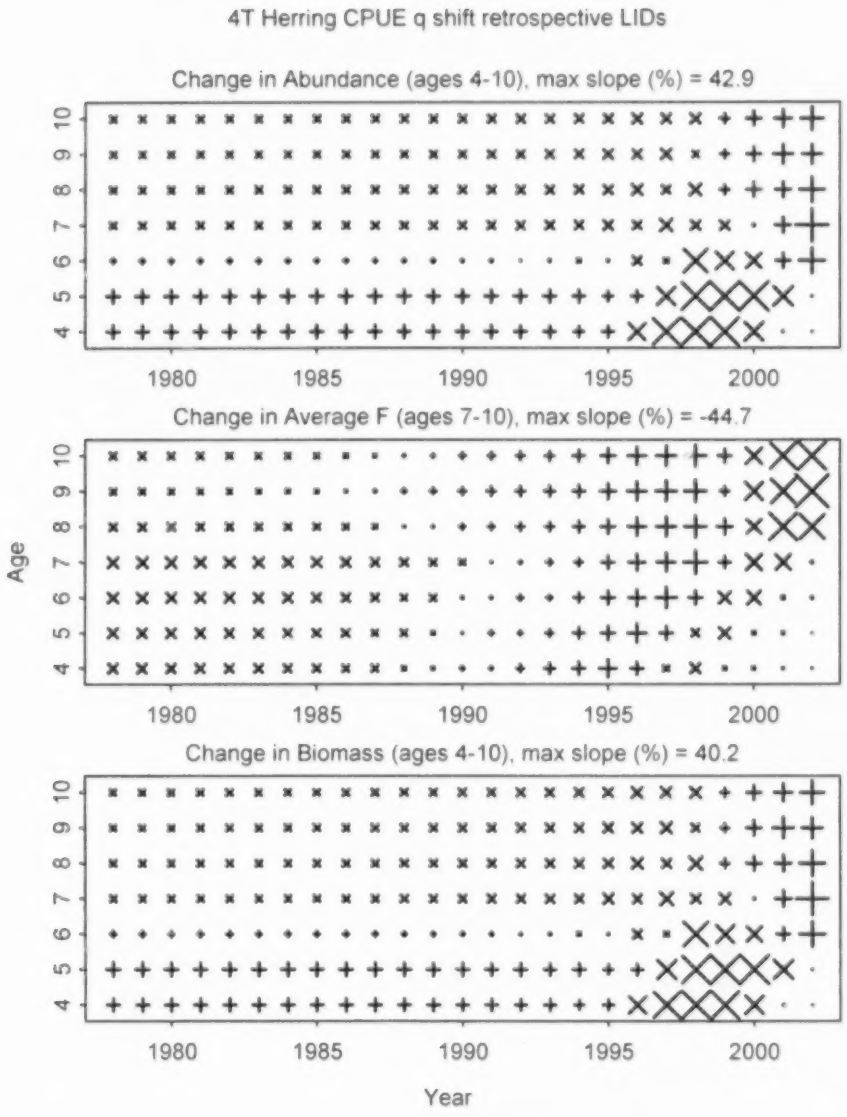

Figure 20: CPUE catchability local influence diagnostics (LID's) for ρ. Each panel shows the elements of $s_{\max}$. The size and type of the plotting symbols are proportional to the absolute value and sign of the elements, respectively. Negative is denoted by an $\times$. Panel a: N_+; Panel b: $\bar{F}$; Panel c: B_+. At the top of each panel $\dot{\rho}_{\max}$ is shown in percent of the ρ values in Figure 2.

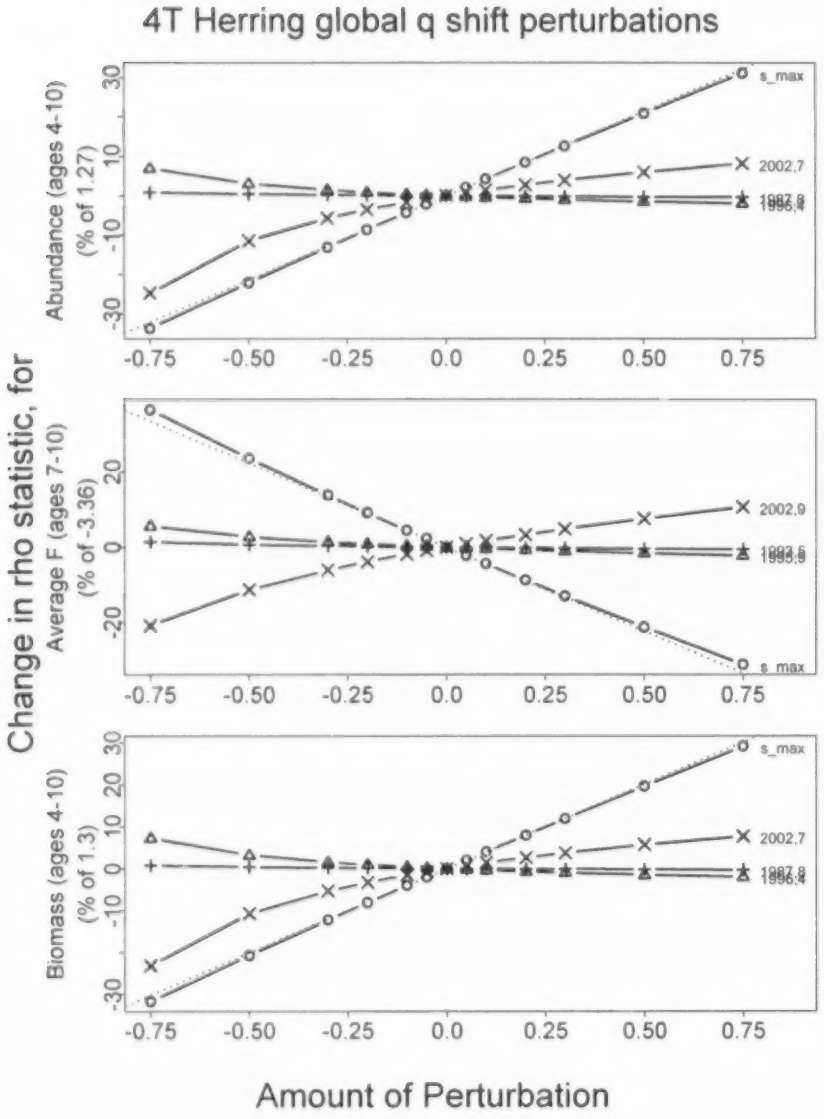

Figure 21: Displacement in ρ's based on some CPUE q global perturbations. All results are in percent of unperturbed estimates. The (year,age) indicates a perturbation of q only for that age and year. The dashed line is a straight line with a slope equal to $\dot{\rho}_{\max}$. Panel a: N_+; Panel b: $\bar{F}$; Panel c: B_+.

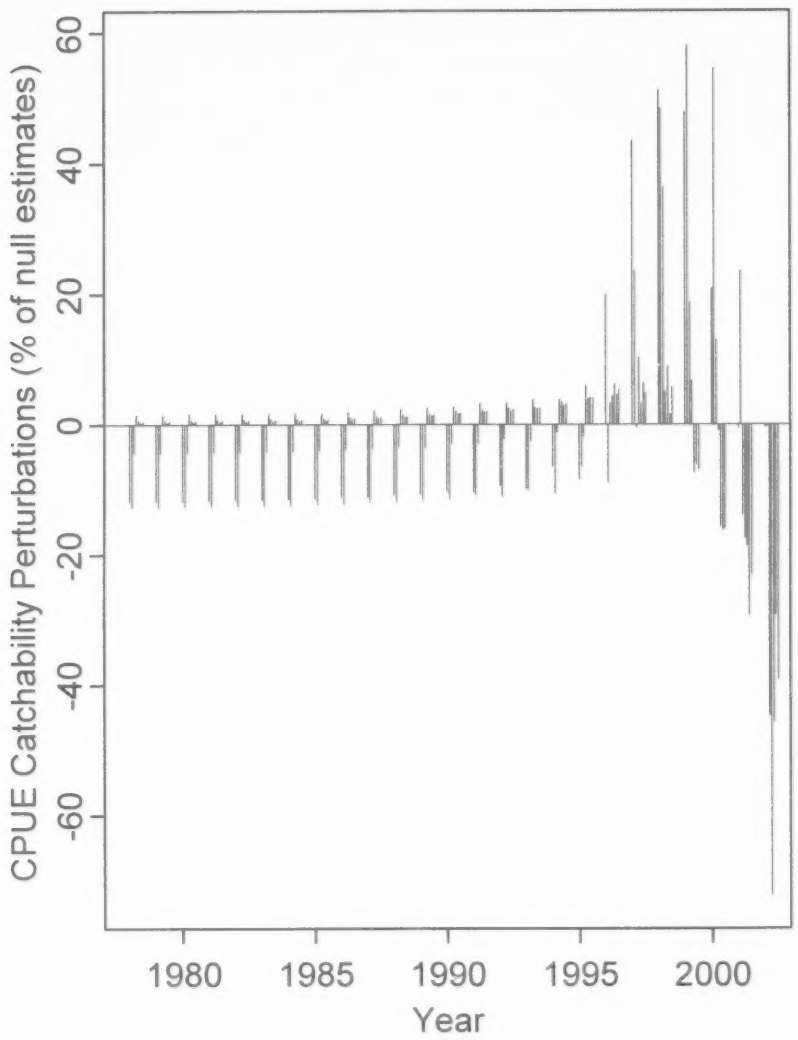

Figure 22: CPUE catchability perturbations to reduce the retrospective pattern in B_+. Each vertical line shows the perturbation to q. Perturbations are clustered by year and shown sequentially for ages 4-10. The values are $-h \times d_{\max}$ in percent with $h = 2$.

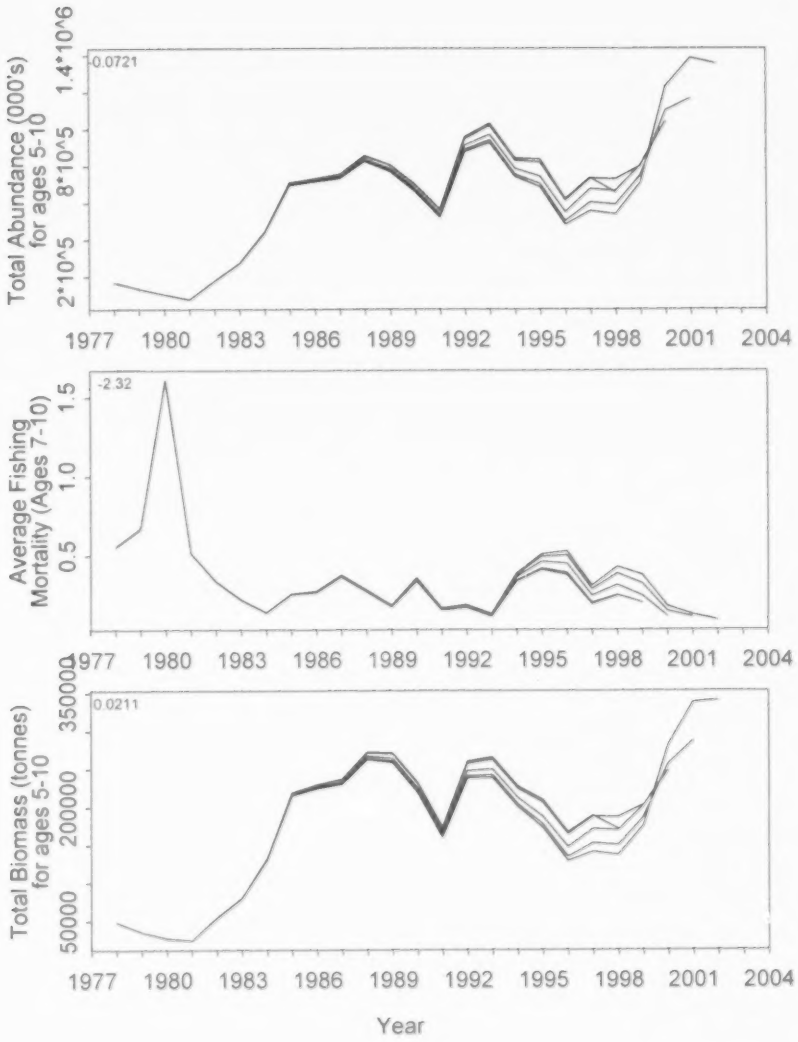

Figure 23: Retrospective estimates based on B_+ $s_{\max}$ perturbed CPUE q's with $h = -2$. The value of ρ is shown in the top left-hand corner. Top panel: N_+; middle panel: $\bar{F}$; bottom panel: B_+.

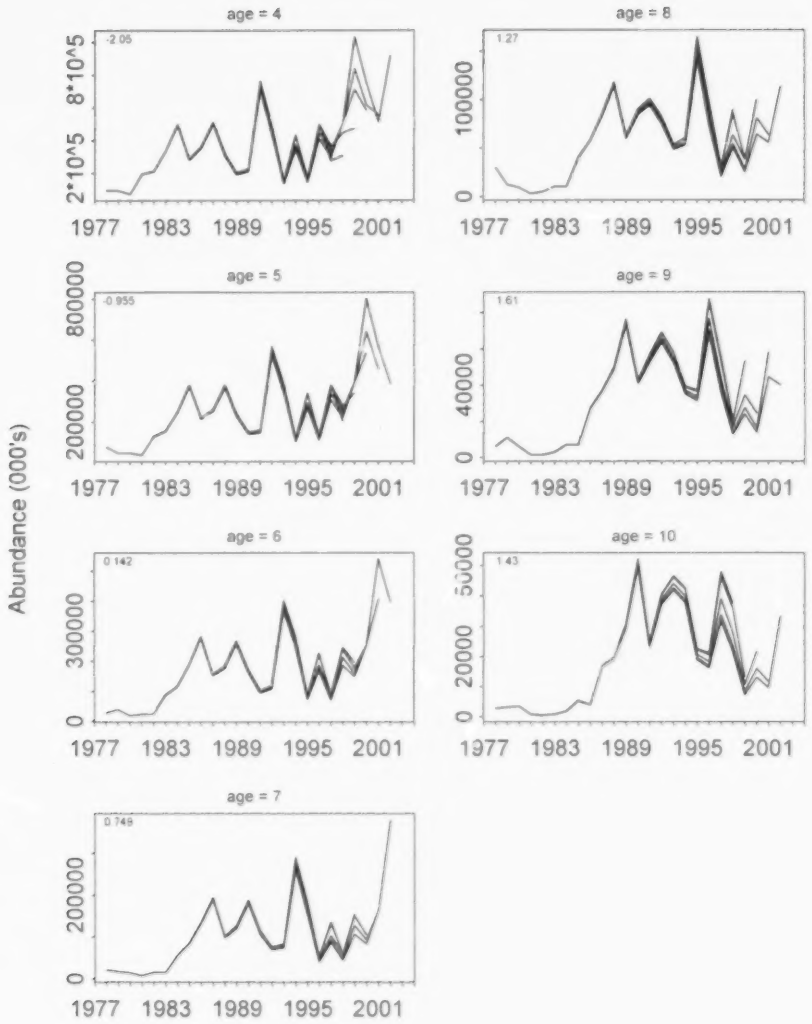

Figure 24: Retrospective estimates of 4T herring SPA stock abundance-at-age estimates based on B_+ $s_{\max}$ perturbed CPUE q's with $h = -2$. Each panel shows the results for an age. The retrospective ρ statistic is shown in the top left-hand corner of each panel.

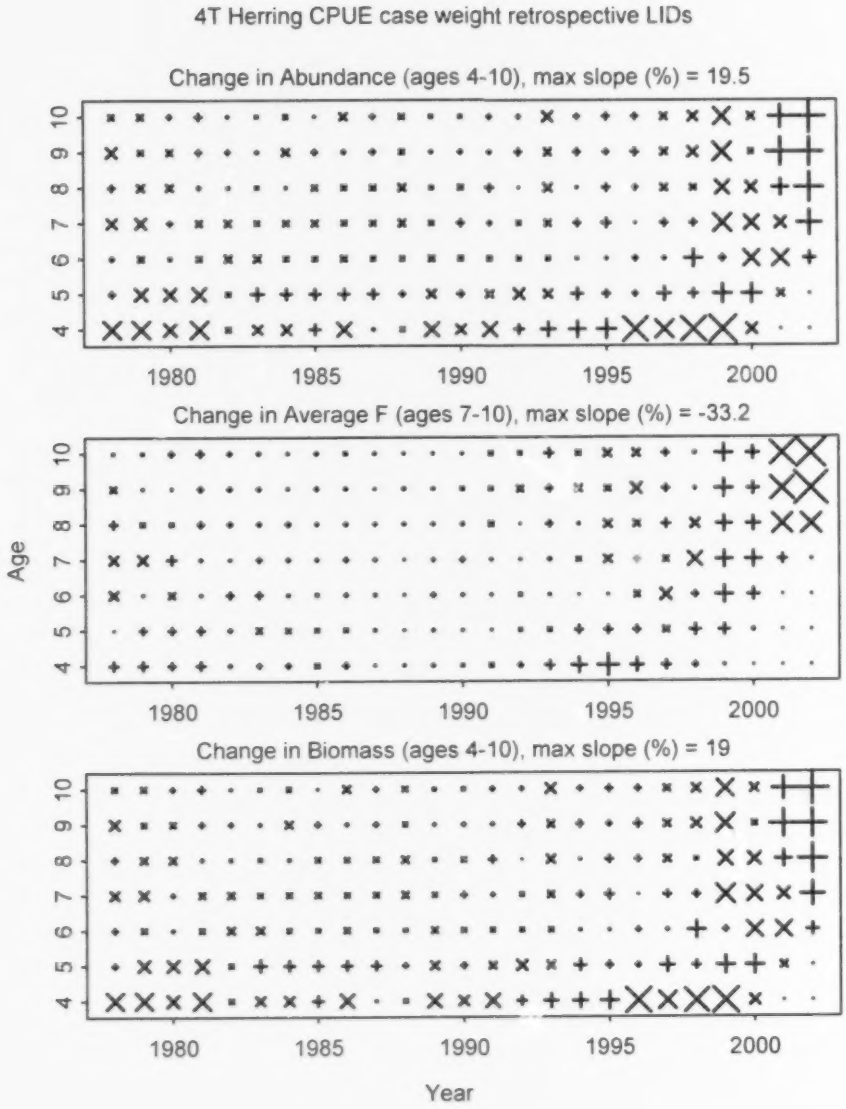

Figure 25: Case weight local influence diagnostics (LID's) for ρ. Each panel shows the elements of $s_{\max}$. The size and type of the plotting symbols are proportional to the absolute value and sign of the elements, respectively. Negative is denoted by an $\times$. Panel a: N_+; Panel b: $\bar{F}$; Panel c: B_+. At the top of each panel $\dot{\rho}_{\max}$ is shown in percent of the ρ values in Figure 2.

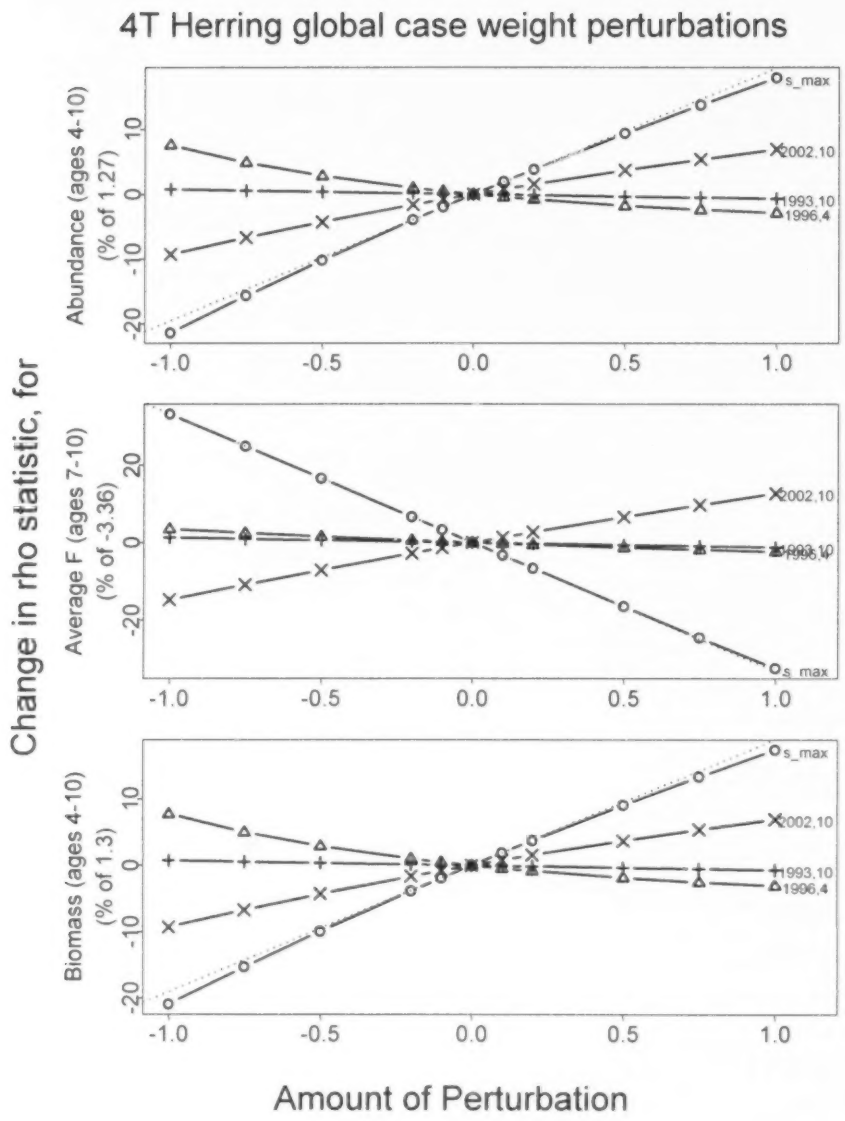

Figure 26: Displacement in ρ's based on some case weight global perturbations. All results are in percent of unperturbed estimates. The (year,age) indicates a perturbation of a case weight for that age and year. The dashed line is a straight line with a slope equal to $\dot{\rho}_{\max}$. Panel a: N_+; Panel b: $\bar{F}$; Panel c: B_+.

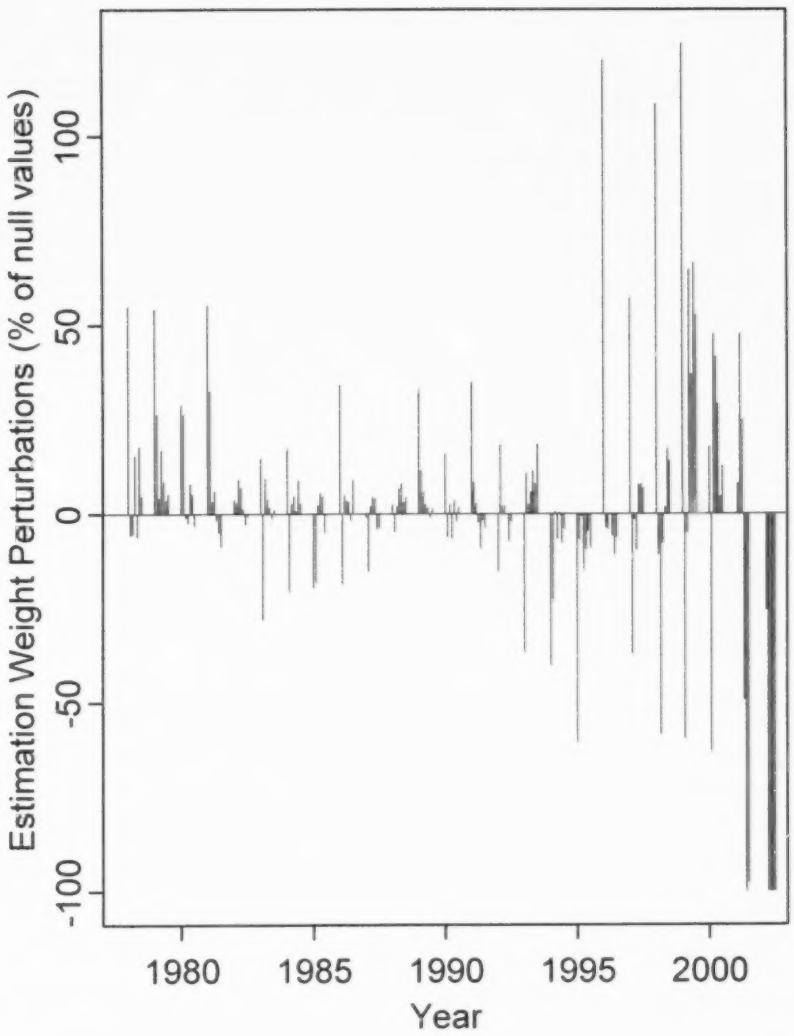

Figure 27: Case weight perturbations to reduce the retrospective pattern in B_+. Each vertical line shows the perturbation to the case weight. Perturbations are clustered by year and shown sequentially for ages 4-10. The values are $-h \times d_{\max}$ in percent with $h = 5$.

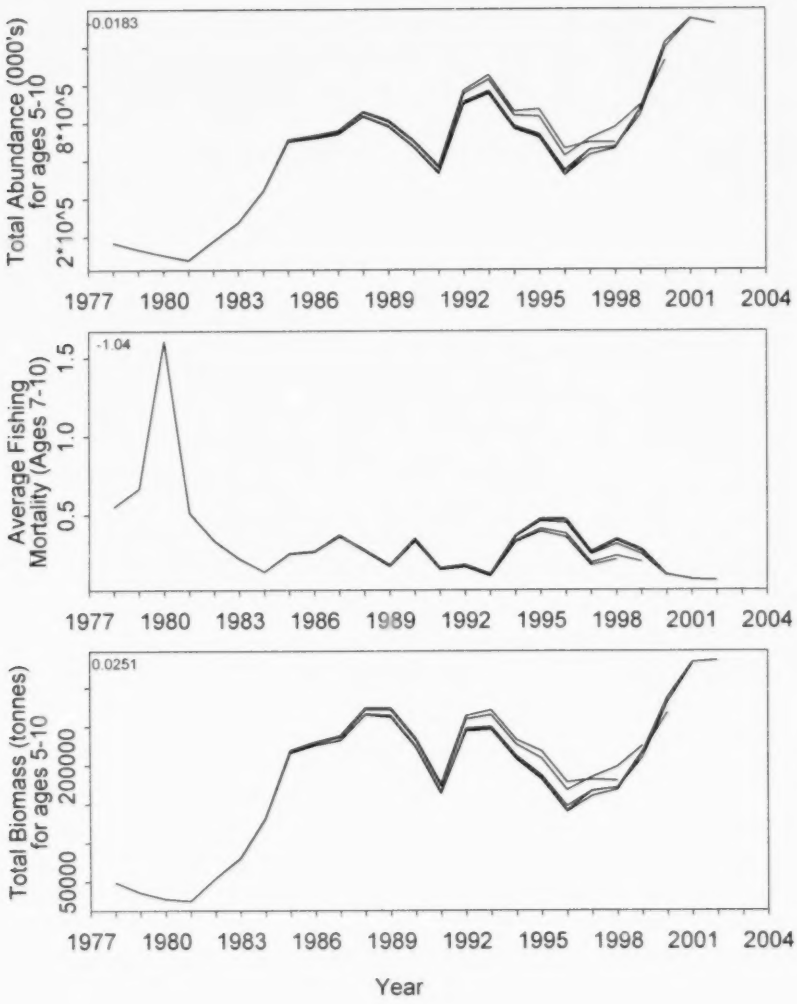

Figure 28: Retrospective estimates based on B_+ $s_{\max}$ perturbed case weights with $h = -5$. The value of ρ is shown in the top left-hand corner. Top panel: N_+; middle panel: $\bar{F}$; bottom panel: B_+.

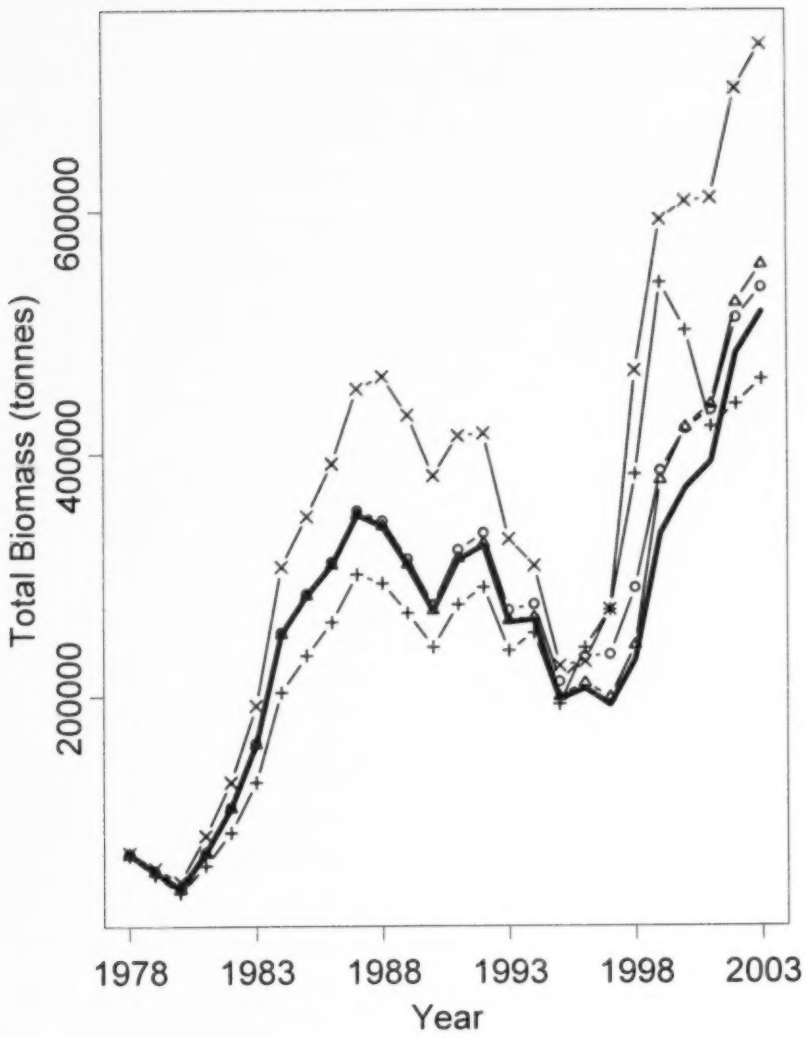

Figure 29: Comparison of total biomass (ages 4-10, B_+) estimates: unperturbed (heavy solid line), M-perturbed (+), catch perturbed ($\times$), q-perturbed ($\triangle$), and case-weight perturbed ($\circ$)